Markov processes and applications to queueing/risk/storage theory and mathematical biology

Florin Avram*

Contents

1 Introduction aux processus stochastiques/aléatoires **3**

2 Marches aléatoires et récurrences **4**
2.1 Prélude: la méthode du conditionnement sur le premier pas 4
2.2 Marches aléatoires 6
2.3 Moments et cumulants 7
2.4 Ruine du joueur pour la marche aléatoire simple 8
2.5 L'operateur associé à la marche simple et ses fonctions harmoniques 10
2.6 Problème de premier passage sur un domaine semi-borné 14
2.7 Equations differentielles et récurrences linéaires 15
2.7.1 L'équation de récurrence linéaire à coefficients constants 15
2.7.2 La méthode des fonctions génératrices 17
2.7.3 Equations differentielles et récurrences linéaires à coefficients polynomiaux 17

3 Problèmes d'absorbtion/Dirichlet pour les chaînes de Markov **18**
3.1 Les chaînes de Markov absorbantes 18
3.2 Les problèmes de Dirichlet 18
3.3 Les espérances des temps d'absorbtion 19
3.4 Les probabilités d'absorbtion 20
3.5 Les distributions des temps d'absorbtion (de type phase/"matrix geométric" 22
3.6 L'opérateur associé à une chaîne de Markov 24
3.7 Une classification des quelques problèmes concernant les processus de Markov 24
3.8 Exercices 25
3.9 Solutions 26
3.10 Projet 31

*Dept. de Math., Université de Pau, France, E-mail: Florin.Avram@univ-Pau.fr

4 Le processus de Poisson **33**
4.1 Rappels sur la distribution de Poisson et exponentielle 33
4.2 Le processus de Poisson multidimensionel 35
4.3 Processus de comptage et renouvellement en temps continu 36
4.4 Le processus de Poisson unidimensionel 37
4.5 Le processsus de Poisson comme limite des processus de Bernoulli 39
4.6 Le processus de Poisson composé 40
4.7 Conclusions 41
4.8 Exercices 41
4.9 La proprietè de Markov 43

5 Les semigroupes (des matrices de transition) de Markov en temps continu **44**
5.1 La propriété de Markov du processus de Poisson et son générateur 46
5.2 Les équations différentielles de Chapman-Kolmogorov 48
5.2.1 Résolution des èquations Chapman-Kolmogorov pour le processus de Markov à deux états 48
5.2.2 Résolution des èquations Chapman-Kolmogorov pour le processus de Poisson 50

6 Semigroups of positive operators **51**
6.1 Stationary/autonomous problems 52
6.2 Time dependent/transient problems 53

7 Les processus de naissance-et-de-mort **55**
7.1 Les files d'attente Markoviennes 56
7.2 Transformées de Laplace des probabilités transitoires de la file M/M/1 57
7.3 Distribution stationnaire et comportement asymptotique 58
7.4 Mesures de performance des files d'attente 61
7.5 La loi distributionelle de Little pour la file M/G/1 (*) 61
7.6 Le problème de Dirichlet/première passage 63
7.7 La file M/M/1 et l'équation de Poisson 64
7.8 Processus quasi-naissance-et-mort/quasi-Toeplitz 65
7.9 Exercices 65

8 Exit/first passage problems **71**
8.1 Actuarial and queueing origins 71
8.2 Levy processes 75
8.3 Wong-Pearson diffusions 77

9 First passage of affine spectrally negative jump-diffusions **78**
9.1 Integro-differential equations for penalty-reward functions of stopped "state dependent Levy processes" 78
9.2 The double Laplace transform of Gerber-Shiu functions 81
9.3 The scale function and two-sided exit problems 82

9.4 Examples of the ODE approach 84
9.4.1 Generalized Pearson process with exponential claims 84
9.5 First passage for Pearson diffusions 87

10 Transient distributions and first-passage probabilities of Markovian processes 87

11 Sheffer and convolution polynomials 89
11.1 Polynomials of convolution and binomial type 89
11.1.1 Characterization by the generating function/sequence 90
11.1.2 Characterization by Δ/lowering operators 90
11.2 Sheffer polynomials 92
11.3 Appell polynomials 92

1 Introduction aux processus stochastiques/aléatoires

Beaucoup de problèmes en physique, biologie, etc, ramène à l'étude des **champs aléatoires**, qui sont des collections des variables aléatoires $X_t, t \in \mathcal{I}$, où l'ensemble des indices $\mathcal{I}$ peut-être $\mathbb{N}^d, \mathbb{Z}^d, \mathbb{R}^d$, un ensemble fini, etc. Pour le cas des indices unidimmensionels $\mathcal{I} = \mathbb{N}(\mathbb{Z})$ et $\mathcal{I} = \mathbb{R}$ on utilise aussi le nom **processus stochastiques** (à temps discret, respectivement continu).

Définition 1.1 *Soit $\mathcal{I}$ un ensemble quelconque. On appelle processus aléatoire X indexé par $\mathcal{I}$ toute famille $(X_t)_{t\in\mathcal{I}}$, de vecteurs aléatoires définis sur un même espace de probabilité $(\Omega, \mathcal{A}, P)$ et à valeurs dans d'états $\mathcal{E}$. Celui la peu-être $\mathcal{E} = \mathbb{R}^p$, $\mathbb{C}^p$, ou même un espace des fonctions comme $\mathcal{E} = C_{[0,\infty)}$, $C^{(p)}_{[0,\infty)}$, etc.*

Note: Lorsque $\mathcal{E} = \mathbb{R}^p$ et $p = 1$, une seule valeur est observée à chaque "instant" t, alors que lorsque $p > 1$, plusieurs variables sont observées et on parle de processus multidimensionnels ou multivariés.

L'espace $\mathcal{I}$ est souvent le temps, ainsi:
$\mathcal{I} = \mathbb{N}$: instants successifs à partir d'un instant initial t_0.
$\mathcal{I} = \mathbb{Z}$: instants successifs avant et après un instant t_0.
$\mathcal{I} = \mathbb{R}$ ou $\mathbb{R}^+$: idem mais processus à temps continu.
$\mathcal{I} = \mathbb{Z}^2$: images.
$\mathcal{I} = \mathbb{Z}^3$: modèle de la matière.

Nous allons considèrer ici surtout des processus à indices unidimmensionels $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ (les premièrs deux cas étant appellés aussi séries chronologiques en statistique). L'étude est facilité alors par l'existence d'un ordre complet entre les indices.

Dans le cas des espaces d'états $\mathcal{E}$ **finis ou dénombrables**, les variables $X_i, i \in \mathcal{I}$ sont appellées discrètes; pour $\mathcal{E} = \mathbb{R}^d$, on parle des variables continues. Le cas discret est le cas plus simple, car il permet d'éviter plusieurs details téchniques (par exemple, dans ce cas, l'ensemble des evenements mesurables pour une variable X_{i_0} est simplement l'ensemble de toutes les parties de $\mathcal{E}$).

Pour modéliser un champs/processus il est necessaire de spécifier de manière consistente l'ensemble de toutes ses **distributions jointes d'ordre fini**.

Définition 1.2 *Soit $X_\cdot = X_t, t \in \mathcal{I}$ un champs aléatoire et soit $J \subset \mathcal{I}$ un sous ensemble fini. On dénotera par X_J la distribution jointe des variables $X_t, t \in J$. L'ensemble $X_J : J \subset \mathcal{I}, |J| < \infty$ sera appellé la famille des distributions jointes d'ordre fini de $X_\cdot$*

En pratique, des proprietés supplementaires sont necessaires pour reduire la complexité inherente dans le modèle ci dessu.

Les processus à variables X_t indépendants sont simples à utiliser, mais peut-être trop simples pour modéliser des phénomènes intéressants. dans ce cas, les distributions jointes sont simplement des produits.

Le prochaîne degré de complexité est donné par les processus Markoviens. Comme en temps discret, les processus de Markov sont beaucoup plus faciles à étudier que les autres processus stochastiques (car leur distributions jointes factorisent), et nous nous restreindront ici à ce cas. Il inclue les marches aléatoires $S_n = \sum_{i=1}^n Z_i$, avec Z_i i.i.d., mais comme ce case est aussi une trasformation simple des processus à variables indépendants, ça simplifie son étude (par exemple la démonstration du CLT).

La classe des processus Markoviens est extrememement riche, avec une complexité qui depend des ensembles $\mathcal{E}, \mathcal{I}$.

2 Marches aléatoires et récurrences

2.1 Prélude: la méthode du conditionnement sur le premier pas

Exercice 2.1 *On lance une monnaie biaisée, avec la probabilité de sortir face égale à q et celle de sortir pile égale à $p = 1 - q$, jusqu'à ce qu'on obtient une pile. Soit $\tilde{N}$ le nombre de faces précedant la première pile ($\tilde{N} = 0, 1, 2, ...$)*

1. *Quelle est la valeur de $\tilde{N}$ si $X_1 ... X_{k-1} = F$ et $X_k = P$?*
2. *Quelle est la loi (distribution) de $\tilde{N}$?*
3. *Trouvez l'espérance $\bar{N}$, par la méthode du conditionnement.*
4. *(*) Trouvez l'espérance $\bar{N}$, par la méthode des fonctions génératrices, i.e: calculer la fonction génératrice des probabilités $p^*(z) = \sum_{k=0}^{\infty} p_k z^k$, en énumérant l'ensemble G de tous les cas possibles, et en observant une relation de recurrence qu'il satisfait. En suite, calculez $p'(1)$.*
5. *Calculez p_k, en developpant la fonction génératrice en somme des puissances.*
6. *(*) Répéter l'exercice antérieur, en prenant pour $\tilde{N}$ le nombre de pas jusqu'à ce qu'on obtient deux piles consécutives.*

Solutions: 2. On a à faire avec une distribution bien connue (la géométrique!!).

3. Il est quand mêmem intéressant de remarquer que l'espérance peut aussi se calculer par un conditionnement sur le premier pas:

$$\bar{N} = p \times 0 + q(1 + \bar{N}) \Longleftrightarrow \bar{N} = \frac{q}{p}$$

Nt: Pour l'autre définition d'une variable g'eométrique $N = \tilde{N} + 1 \in \{1, 2, ...\}$ (en incluant la première face), on obtient directement $EN = E(\tilde{N} + 1) = \frac{1}{p}$.

On peut aussi conditionner sur le premier pas:

$$\begin{aligned} n &= \mathbb{E}[N|X_0 = 1] = \mathbb{P}[X_1 = 2]\mathbb{E}[N|\{X_1 = 2, X_0 = 1\}] + \mathbb{P}[X_1 = 1]\mathbb{E}[N|\{X_1 = 1, X_0 = 1\}] \\ &= \mathbb{P}[X_1 = 2]1 + \mathbb{P}[X_1 = 1](1 + \mathbb{E}[N|\{X_0 = 1\}]) = p * 1 + q * (1 + n) = 1 + q * n \qquad (1) \end{aligned}$$

Note: La deuxième ligne est une conséquence de la décomposition "premier pas + le reste"

$$N = 1 + N'$$

Après ça, on utilise le fait que les distributions conditionnées par départ du "reste" $N'|X_1 = i$ sont connues:

1. $(N'|X_1 = 2) \equiv 0$ et
2. $$\boxed{\mathcal{L}(N'|X_1 = 1) = \mathcal{L}(N|X_0 = 1)}$$

   par la proprieté de Markov=oubli du passé, et par le fait que la seule difference entre les réalisations possibles de N' et ceux du N et le moment "de départ de la montre", qui est indifférent pour une chaîne avec matrice de transition stationnaire.

4. $$G = \{H, TH, TTH, TTTH, ...\} = \{H\} \cup TG \qquad (2)$$

Remplaçons chaque terme par $p^{nb.H} q^{nb.T} z^{nb.H+nb.T}$, et ajoutons tous les termes. Observons qu'en rajoutant tous les termes avec $nb.H + nb.T = n$, on obtient precisement $p_n z^n$, et donc la somme totale est la fonction génératrice $p(z) = \sum_n p_n z^n$. De l'autre coté, la décomposition (2) implique $p(z) = p + qzp(z)$, et $p(z) = \frac{p}{1-qz}$. Finalement, $\bar{N} = p'(1) = \frac{q}{p}$.

$$p^*(z) = \frac{p}{1 - qz} = p \sum_{k=0}^{\infty} q^k z^k$$

5. En developpant $p_n = pq^n, n = 0, 1, ...$, on trouve la distribution d'une variable g'eométrique (plus grande que 0).

Exercice 2.2 *Un spéolog est perdu dans une grotte, avec deux sorties U, O, les chemins de la quelle forment le graphe papillon ci-dessous. Il se trouve à present dans le point A, et à cause de la manque de visibilité, est obligé a faire une* **marche aléatoire** *entre les sommets du papillon.*

a) Calculez l'espérance du temps qu'il prendra à sortir, en sachant que les chemins menant à U prennent deux heures, ceux menant à O trois heures, et les autres 5 heures.

b) La grotte contient aussi un génie, qui frappe le spéolog des nb. aléatoires des coups, avec esperance 3 sur les chemins menant à U, 5 pour ceux menant à O, et 2 pour les autres chemins. Calculez l'espérance du nombre des coups que le spéolog prendra avant de sortir.

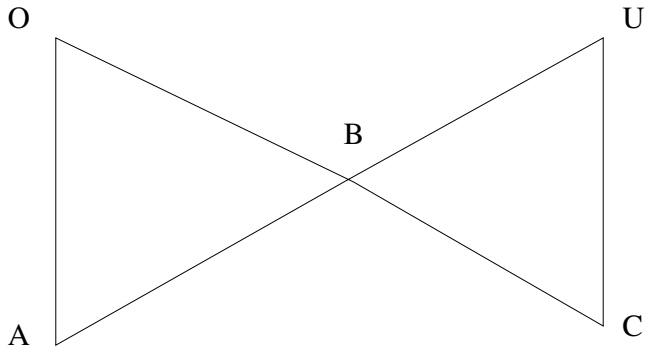


Figure 1: **Marche aléatoire simple sur le graphe papillon**

2.2 Marches aléatoires

Définition 2.1 Marches aléatoires

Soit $Z = (Z_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires **i.i.d** *(i.e. indépendantes et de même loi), à valeurs dans un groupe G, et soit $X_0 \in G$ indépendant de Z.*

Le processus $X_n \in G, n = 0, 1, ...$ donné par la somme de ces variables

$$X_n = X_0 + Z_1 + Z_2 + \cdots + Z_n, \quad n \in \mathbb{N} \tag{3}$$

s'appelle **marche aléatoire**. *Comme alternative, la marche aléatoire peut-être definie récursivement par la récurrence*

$$X_n = X_{n-1} + Z_n, n = 1, ... \tag{4}$$

Motivation: Les marches aléatoires sont parmi les processus stochastiques les plus utiles (par exemple en physique, mathématiques financières, files d'attente, statistique, etc...). Ils peuvent servir comme des excellents exemples introductifs, qui motiveront un étude plus approfondi des processus stochastiques. Finalement, ils sont aussi parmi les processus les meilleurs compris, car ils ramènent souvent à des solutions analytiques, ce qui rendent les résultats plus transparents.

Exemple: Les marches aléatoires sur $\mathbb{Z}$

Définition 2.2 Marches aléatoires sur $\mathbb{Z}$.

Soit $Z = (Z_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires **i.i.d** *(i.e. indépendantes et de même loi), à valeurs en $\mathbb{Z}$, ayant une distribution discrète $\boldsymbol{p} = (p_i, i \in \mathbb{Z})$, et soit $X_0 \in \mathbb{Z}$ indépendant de Z et ayant aussi une distribution discrète.*

a) Le processus $X_n \in \mathbb{Z}, n = 0, 1, ...$ donné par la somme de ces variables

$$X_n = X_0 + Z_1 + Z_2 + \cdots + Z_n, \quad n \in \mathbb{N} \tag{5}$$

s'appelle **marche aléatoire**. *Comme alternative, la marche aléatoire peut-être definie récursivement par la récurrence*

$$X_n = X_{n-1} + Z_n, n = 1, ... \tag{6}$$

b) Si en plus $|Z_n| = 1$, i.e. $p_i \neq 0$ ssi $i = \pm 1$, et le processus (5) est appelé **marche aléatoire simple**. *On denotera p_1 par p et de lors la distribution de Z_n est de la forme $p\delta_1 + (1-p)\,\delta_{-1}$, i.e. $P[Z_n = 1] = p$ et $P[Z_n = -1] = 1 - p$ avec $0 < p < 1$.*

Si $p = q = .5$ on parle d'une **marche aléatoire symmetrique**, *et avec $p \neq q$ on parle d'une* **marche aléatoire biaisée**.

2.3 Moments et cumulants

Exercice 2.3 Les moments et cumulants de la marche simple. *Soit $X_0 = 0 \in \mathbb{N}$ le capital initial d'un joueur. Au temps $n = 1, 2, ...$, le joueur gagnera $Z_n = 1$ avec probabilité p et perdera $Z_n = -1$ avec probabilité $1 - p$, où $0 < p < 1$. Soit $X_n = X_0 + Z_1 + Z_2 + \cdots + Z_n$ son capital au temps n.*

Calculez:

1. *L'esperance du son gain $e_n = EX_n$.*
2. *La variance du son gain $v_n = \operatorname{Var} X_n$.*
3. *La fonction génératrice des moments $M(u, n) = Ee^{uX_n}$.*
4. *La fonction génératrice des cumulants $\kappa(u, n) = \log(Ee^{uX_n})$.*

Notes: 1) Il est clair que ces propriétés de linéarité (de l'espérance , de la variance, et de la fonction génératrice des cumulants), sont vraies pour chaque marche aléatoire.

2) La connaissance de la distribution ou de la fonction génératrice des moments d'une variable X sont typiquement equivalents. Mais, pour une somme $\sum_{i=1}^n Z_i$ des v.a. i.i.d., pendant que la comme distribution est la n-ième convolution $\boldsymbol{p}^{*,n}$ de la $\boldsymbol{p}$ distribution de Z_i, la fonction génératrice des moments $Ee^{\theta\sum_{i=1}^n Z_i}$ est beaucoup plus simple à obtenir (èatnt la n-ième puissance de la fonction génératrice des moments $Ee^{\theta Z_1}$).

Exercice 2.4 *Soit $m_n, n = 0, 1, 2, ...$ les moments d'une va X, soit $\kappa_X(u) = \log M_X(u) = \log(\sum_n \frac{u^n}{n!} m_n) = \sum_n \frac{u^n}{n!} c_X(n)$ la fonction génératrice des cumulants, et soit $C_X(n) = \frac{\partial^n \kappa(u)}{(\partial u)^n}_{u=0}$ les cumulants.*

a) Montrez, en utilisant un logiciel symbolyque, que $\forall X, c_x(0) = 0, c_X(1) = m_1, c_X(2) = \mathrm{Var}\,(X), c_X(3) = m_3 - 3m_1m_3 + 2m_1^3$.

b) Montrez, en utilisant un logiciel symbolyque, que $\forall X, m_2 = c(2) + c_1^2, m_3 = c(1)^3 + 3c(1)c(2) + c_3$.

Nt: 1) Le cumulant d'un ordre donné est un polynome dans les moments d'ordre plus petit ou egale, et reciproquement.

2) Les coefficients de l'expansion des moments en fonction des cumulants sont donné par des nombres des partitions.

3) Les cumulants d'une variable centré ($m_1 = 0$) coincide avec les moments jusqu'au troisième ordre. C'est le quatrième cumulant, la "kurtosis", donné dans le cas centré par $c(4) = m_4 - 3m_2^2$, qui joue un role important dans certaines tests statistiques (comme de nonnormalité, par exemple).

Exercice 2.5 *Demontrer le théorème de limite centrale:*

$$\lim_{n\to\infty}^{(d)} n^{-1/2} \sum_{i=1}^{n} (Z_i - EZ_i) = N_{(0,\mathrm{Var}\,Z)}$$

où $\lim^{(d)}$ *denote la convergence des distributions, en utilisant:*

1. *le fait que la convergence des distributions est equivalente a celle des fonctions génératrice des moments, et*
2. *le fait que* $Ee^{uN_{0,v}} = e^{vu^2/2}$*, i.e. la variable normale a tous les cumulants d'ordre plus grand que trois nuls.*

Exercice 2.6 *Pour la marche simple, calculez*

1. *Le premier, deuxième et troisième cumulants* $\kappa_i(n), i = 1, 2, 3$ *de* X_n*, i.e. les premiers trois coefficients dans l'expansion* $\kappa(u, n) = \sum_i \kappa_i(n)u^i$ *en puissances de* u*.*
2. *Le deuxième moment de* X_n*. Quelle est la particularité du cas* $p = 1/2$*?*
3. *Le troisième moment de* X_n*.*

2.4 Ruine du joueur pour la marche aléatoire simple

Nous étudiérons maintenant un problème concernant les probabilités d'absorbtion dans un ensemble d'arrêt et d'autres problèmes similaires, les solutions des quelles sont disponibles explicitement dans le cas de la marche aléatoire simple unidimensionnelle.

Exemple 2.1 La ruine du joueur et autres "problèmes de Dirichlet" pour la marche aléatoire simple. *Considérons la marche aléatoire simple*

$$X_n = X_0 + Z_1 + Z_2 + \cdots + Z_n, X_n \in \mathbb{Z}$$

avec (Z_n) une suite de variables aléatoires réelles indépendantes de même loi $P[Z_n = \pm 1] = p, q$. Nous étudierons la marche jusqu'au "temps d'arrêt/sortie" $T = \min[T_0, T_B]$ quand le process sort de l'interval $[0, B]$ pour B donné, i.e. on prend 0 et B comme états absorbants. On appelle ce problème **la ruine du joueur***, a cause de l'interpretation d'un joueur qui s'engage dans une série de parties (indépendantes) à un jeu où à l'issue de chaque partie il gagne 1F avec une probabilité p et perd 1F avec une probabilité $q = 1 - p$, et qui décide de s'arrêter de jouer dès qu'il aura B francs en poche, ou dès qu'il n'a plus d'argent. Pour tout $n \in \mathbb{N}$, on note X_n la fortune du joueur au bout de n parties, et $X_0 = i$ représente sa fortune à l'entrée dans le Casino. On dénotera par $\mathbb{E}_i$ l'esperance en commençant de i (conditionnant sur $X_0 = i$), et on designe par E l'événement que le joueur gagne, i.e.*

$E = \{x_T = B\} = [\exists n \in \mathbb{N}$ *tel que* $X_n = B, X_i > 0, i = 1, ..., n-1]$.

Pour tout i de $\{0, ..., B\}$, on pose :

$$b_i = P(E \mid [X_0 = i]).$$

1. *Calculer b_0 et b_B .*

2. *Montrer que :*

$$\forall\, i \in \{1, ..., B-1\} \ , \ b_i = p\, b_{i+1} + q\, b_{i-1} \ \textit{(on rappelle que } q = 1 - p\textit{).}$$

3. *Obtener une expression explicite de b_i pour tout i de $\{1, ..., B\}$.* Indication: *Remarquez que la solution satisfaisant $b_0 = 0$ est de la forme:*

$$b_i = \begin{cases} k\,(1 - (\frac{q}{p})^i) & \textit{quand } p \neq q \\ k\, i & \textit{quand } p = q \end{cases}$$

et déterminer k tq la condition frontière de b_B soit satisfaite.

4. *Pour tout i de $\{0, ..., B\}$, on pose $a_i = P(F \mid [X_0 = i])$ où F est l'événement "le joueur repart ruiné" . En procédant comme auparavant, montrer que :*

$$a_i = \begin{cases} \frac{\left(\frac{q}{p}\right)^i - \left(\frac{q}{p}\right)^B}{1 - \left(\frac{q}{p}\right)^B} \ \textit{si } p \neq \frac{1}{2} \\ \\ \frac{B-i}{B} \ \textit{si } p = \frac{1}{2} \end{cases}$$

Pour tout i de $\{0, ..., B\}$, calculer $a_i + b_i$. Que peut-on en déduire?

Calculez les probabilités de ruine quand $B \to \infty$, pour $p > q$ et pour $p \leq q$. Expliquez la rélation avec le comportement de $X_t, t \to \infty$.

5. *Obtenez un système d'équations pour l'espérance du gain final $f_i = \mathbb{E}_i X_T$. Calculez cette fonction pour $p = q$.*

6. *Obtenez un système d'équations pour l'espérance du temps de jeu: $t_i = \mathbb{E}_i T$. Calculez cette fonction, pour $p = q$, et pour $p < q$, quand $B \to \infty$.*

7. *Obtenez un système d'équations pour l'espérance du "coût cumulé d'inventoire"* $c_i = \mathbb{E}_i \sum_{t=0}^{T-1} X_t$. *Calculez cette fonction, pour* $p = q$, *et pour* $p < q$, *quand* $B \to \infty$.

8. *Obtenez un système d'équations pour* $w_i = \mathbb{E}_i a^T = \sum_{k=0} \mathbb{P}_i[T = k] a^k$ *(qui est la fonction génératrice des probabilités* $\mathbb{P}_i[T = k]$*). Calculez cette fonction, pour* $p \neq q$.

9. *Obtenez les équations de récurrence et les conditions frontière satisfaites par* $u_x = \mathbb{E}_x a^T g(X_T), a \in (0,1)$ *et par* $v_x = \mathbb{E}_x[a^T g(X_T) + \sum_{t=0}^{T-1} h(X_t)], a \in (0,1)$.

On resoudra ceci en utilisant la méthode du **conditionnement** sur le premier pas Z_1, l'idée de quelle est d'obtenir des relations de récurrence qui lient les valeurs de l'espérance conditionnée à partir de tous les points de départ possibles.

2.5 L'operateur associé à la marche simple et ses fonctions harmoniques

Nous verrons, en examinant les questions 2)-8) de cet exercice, qu'ils utilise toutes le même opérateur

$$(Gf)_n := (P - I)(f)_n = p\, f_{n+1} + q\, f_{n-1} - f_n \tag{7}$$

la seule difference étant dans les conditions frontière et dans la partie nonhomogène. En plus, ils se regrouperont en deux types de questions:

1. "Gain final esperé", satisfaisant:

$$f_n = \mathbb{E}_n[g(X_T)] = pf_{n+1} + qf_{n-1} \Longleftrightarrow (Gf)_n = 0, \quad F(0) = g(0), F(B) = g(B)$$

2. "Coût total accumulé esperé"

$$f_n = \mathbb{E}_n[\sum_{0}^{T-1} h(X_i)] = h(n) + pf_{n+1} + qf_{n-1} \Longleftrightarrow (Gf)_n = 0, \quad f(0) = 0, f(B) = 0$$

Solution:

1. $b_0 = 0, b_B = 1$

2. Gain final esperé, $g(x) = 1_{x=B}$.

   En conditionnant, on trouve:

$$\begin{aligned} b_n &= \mathbb{P}_n[X(T) = B] \\ &= p\, \mathbb{P}_n[X(T) = B/X(1) = n+1] + q\, \mathbb{P}_n[X(T) = B/X(1) = n-1] \\ &= p\, b_{n+1} + q\, {}_{n-1} \quad 1 \leq n \leq B-1 \end{aligned}$$

   car

$$\begin{aligned} \mathbb{P}_n[X(T) = B/X(1) = n \pm 1] = \mathbb{P}[X(T) = B/X(0) = n, X(1) = n \pm 1] = \\ \mathbb{P}[X(T) = B/X(1) = n \pm 1] = \mathbb{P}[X(T) = B/X(0) = n \pm 1] = b_{n\pm 1} \end{aligned}$$

   en utilisant la proprieté de Markov et l'homogeneité.

3. Quand $p = q = 1/2$, $b_x = P_x[X(T) = B]$ satisfait:

$$\begin{aligned} b_n &= \frac{b_{n+1}}{2} + \frac{b_{n-1}}{2} \quad \text{for any } 1 \le n \le B-1 \\ b_B &= 1 \\ b_0 &= 0 \end{aligned}$$

La méthode de résolution des équations de récurrence homogènes à coefficients constants commence en cherchant des solutions de la forme $b_n = r^n$. Si les racines de l'équation auxiliaire sont distinctes, la solution générale est:

$$b_n = k_1 r_1^n + k_2 r_2^n$$

où k_1, k_2 sont déterminés en utilisant les conditions frontière.

Ici, cherchant des solutions puissances r^x ramène à l'équation $r^2 - 2r + 1 = 0$ à deux racines identiques $r_{1,2} = 1$. La solution générale est $b_x = A + Bx$. Les conditions frontière donnent $b_x = \frac{x}{B}$.

Solution finale si $p \neq q$: $b_n = \frac{1-(q/p)^n}{1-(q/p)^B}$.

4. $a_i + b_i = 1$, et donc la marche sera eventuellement absorbé dans une des deux frontiéres (elle ne peut pas rester à l'intérieur indéfinimment).

Pour $p = q$, $\lim_{B\to\infty} a_n = \lim_{B\to\infty} \frac{B-n}{B} = 1$. Autrement,

$$\lim_{B\to\infty} a_n = \lim_{B\to\infty} \frac{(q/p)^n - (q/p)^B}{1-(q/p)^B} = \begin{cases} (q/p)^n, & q < p \\ 1, q > p. \end{cases}$$

5. $f_x = E_x[X(T)]$ (valeur finale ésperée) satisfait $Gf(x) = 0, f(0) = 0, f(B) = B$. Pour $p = q$, la solution $f_x = x$ est obtenue comme ci-dessus:

$$\begin{aligned} f_x &= \frac{f_{x+1}}{2} + \frac{f_{x-1}}{2} \quad \text{for any } 1 \le x \le B-1 \\ f_B &= B \\ f_0 &= 0 \end{aligned}$$

(C'est aussi une fonction "harmonique", mais avec conditions frontière différentes.)

6. $t_x = E_x[T]$ (temps de sortie ésperé) est un coût total accumulé esperé (obtenu en prenant $h(x) = 1$), qui satisfait le système inhomogène $Gt(x) + 1 = 0, t(0) = 0, t(B) = 0$.

Pour $p = q$

$$\begin{aligned} t_x &= \frac{t_{x+1}}{2} + \frac{t_{x-1}}{2} + 1 \quad \text{for any } 1 \le x \le B-1 \\ t_B &= 0 \\ t_0 &= 0 \end{aligned}$$

La solution d'une équation nonhomogène est donnée par

$$t_x = t_p(x) + h(x)$$

où $t_p(x)$ est une solution particulière et $h(x)$ est la solution générale de l'équation homogène. Commençons par l'équation homogène.

La solution générale homogène ("fonction harmonique") $h(x) = A + Bx$ pour cet opérateur a été déjà obtenue ci-dessus.

Nous aimerions maintenant trouver une solution particulière $t_p(x)$ de l'équation $Gt_p(x) = -1$ de la même forme que la partie nonhomogène -1 de l'équation, donc $t_p(x) = C$; mais, comme les constantes, et puis aussi les fonctions linéaires vérifient l'équation homogène $Gt_p(x) = 0$, nous devrons modifier deux fois cette forme en multipliant par x, en arrivant donc à $t_(x) = Cx^2$. Comme $Gx^2 = 2x(p-q) + 1 = 1$, on trouve $C = -1$ et finalement la solution particulière $t_p(x) = -x^2$.

La solution générale est donc $t(x) = -x^2 + A + Bx$ et les conditions frontière ramènent à $t_x = x(B - x)$.

Pour $p \neq q$

$$\begin{aligned} t_x &= pt_{x+1} + qt_{x-1} + 1 \quad \text{for any } 1 \leq x \leq B-1 \\ t_B &= 0 \\ t_0 &= 0 \end{aligned}$$

La solution generale homogène avec $p \neq q$ est $h(x) = k_1(q/p)^n + k_2$ et le terme nonhomogène 1 sugere une solution particulière constante k, mais comme ça satisfait l'équation homogène, on modifie à kn. Finalement, $k = \frac{1}{q-p}$.

La solution particulière est $t_p(x) = \frac{x}{q-p}$; elle satisfait deja $t_p(0) = 0$. La partie homogène $h(x) = t_x - t_p(x)$ devra aussi satisfaire $h(0) = 0$ et donc elle sera de la forme $h(x) = A\tilde{h}(x)$ où $\tilde{h}(x) = ((q/p)^x - 1)$.

En demandant que $t_n = \frac{n}{q-p} + A(q/p)^n - 1)$ satisfait la condition frontière $t_B = 0$ on trouve:

$$t_n = t_p(n) - t_p(B)\frac{\tilde{h}(n)}{\tilde{h}(B)} = \frac{n}{q-p} - \frac{B}{q-p}\frac{(q/p)^n - 1}{(q/p)^B - 1}.$$

La limite quand $B \to \infty$ est $t_n = \begin{cases} \infty & \text{si } p > q \\ t_p(n) = \frac{n}{q-p} & \text{si } p < q \end{cases}$; on peut aussi obtenir ce resultat en utilisant l'approximation détérmiste $X_n - X_0 \sim nE(Z_1)$, appellée aussi limite fluide.

7. $c_x = E_x[\sum_0^{T-1} X(t)]$ (coût total d'inventaire ésperé) satisfait le système inhomogène $Gc(x) + x = 0, c(0) = 0, c(B) = 0$.

Pour $p = q$:

$$\begin{aligned} c_x &= \frac{c_{x+1}}{2} + \frac{c_{x-1}}{2} + x \quad \text{for any } 1 \leq x \leq B-1 \\ c_B &= 0 \\ c_0 &= 0 \end{aligned}$$

Une solution particulière est $c_p(x) = \frac{-x^3}{3}$. Finalement, on arrive à $c(x) = \frac{x(B^2-x^2)}{3}$.

Pour $p \neq q$, une solution particulière est $c_p(x) = \frac{x^2}{2(q-p)}$ (elle satisfait deja $c_p(0) = 0$). La partie homogène satisfaisant $h(0) = 0$ sera toujours $h(x) = A\tilde{h}(x)$ où $\tilde{h}(x) = ((q/p)^x - 1)$.

En demandant que $c_n = c_p(n) + A(q/p)^n - 1)$ satisfait la condition frontière $c_B = 0$ on trouve:

$$c_n = c_p(n) - c_p(B)\frac{\tilde{h}(n)}{\tilde{h}(B)}$$

La limite quand $B \to \infty$ est $c_n = \begin{cases} \infty \text{ si } p > q \\ c_p(n) \text{ si } p < q \end{cases}$.

8. On arrive a $w(x) = A_1 z_1^x + A_2 z_2^x$, où z_i sont les racines de $pz^2 - a^{-1}z + q = 0$, et A_i satisfont $A_1 z_1^B + A_2 z_2^B = 1, A_1 + A_2 = 1$ et $w(x) = \frac{z_1^x - z_2^x + z_1^x z_2^x (z_1^{B-x} - z_2^{B-x})}{z_1^Z - z_2^B}$

9. On a $u_x = g(x)$, pour $x \in \{0, B\}$, et le conditionnement donne la relation: $u_x = \mathbb{E}_x[a^T g(X_\tau)] = a(pu_{x+1} + qu_{x-1})$.

   $v_x = g(x)$, pour $x \in \{0, B\}$, et le conditionnement donne la relation: $v_x = a(pv_{x+1} + qv_{x-1}) + h(x)$.

Nous avons vue ici une des idées les plus importantes de la modélisation Markovienne: **les ésperances, vues comme fonctions de l'état initial, satisfont certaines équations qui font toujours intervenir un opérateur associé fixe, appelé générateur du processus, même que les conditions frontière, termes non-homogènes, et d'autre "details" (comme la presence/absence d'une multiple de l'operateur identité) peuvent varier**.

Il s'avère que les mêmes èquations décrivent la solution des problèmes analogues pour toutes les chaîne de Markov à espace d'états comptable, et avec des états absorbants – voir la prochaîne section.

Par exemple, pour les chaînes de Markov, l'operateur associé est $G = P - I$, où P est la matrice de transition, et pour le cas particulier d'une marche aléatoire $X_t = \sum_{i=1}^t Z_i$ avec $p_k = P[Z_i = k], k \in [-c, d]$ on a encore $G = P - I$, où $P = \sum_k p_k F^k$ et F est l'operateur de translation $(Ff)_k = f_{k+1}, k \in \mathbb{Z}$.

Alors, nous obtendrons des èquations similaires pour les problèmes respectives, juste en remplaçant l'ancien operateur par le nouveau.

On recontre la même situation pour toute la classe des processus "de Markov", X_t, différents qu'elles soient, vivant sur des espaces S considerablement plus compliqués, la seule difference étant que l'operateur $G_X : F(S) -> F(S)$ associé a ces processus sera plus compliqué!

Par exemple, les problèmes de cette section ont aussi des versions à espace d'états continu, obtenu en considérant des marches avec incréments infinitésimaux ϵ, et en prenant la limite $\epsilon \to 0$. La marche aléatoire devient ainsi un processus avec chemins continus, appelé mouvement Brownien. Les équations resterons les mêmes, seul l'operateur G changera (dans un operateur differentiel).

En conclusions, il existe une correspondance un á un entre les processus de Markov et une certaine classe des operateurs deterministes associés; nous l'appellerons "Le Dictionnaire".

2.6 Problème de premier passage sur un domaine semi-borné

Soit $\psi_n := \lim_{B\to\infty} a_n(B)$ la probabilité de ruine sur $[0, \infty)$. Nous avons deja vu, en partant des recurrences sur un domaine borné $[0, B]$, que:

$$\psi_n = \begin{cases} (q/p)^n, & q < p \\ 1, & q \geq p. \end{cases}$$

Cette solution peut être trouvée aussi directement, sans passer par la probabilité de ruine sur $[0, B]$, en s'appuyant sur des des arguments probabilistes. On remarque d'abord que cette fonction est multplicative en n, i.e. $\psi_n = \rho^n$, et ensuite on choisit ρ selon la limite de X_n quand $n \to \infty$.

La solution la plus simple est en effet de remarquer que l'absence des sauts strictement plus grands que 1 implique une propriété de multiplicativité, ce qui impose une solution puissance. Mais, bien-sur, cette approche ne peut pas contribuer à l'étude des marches "non-simples" dans les deux directions.

Examinons maintenant la méthode de fonctions génératrices (analogues à la transformée de Laplace), qui n'est pas réellement necessaire pour la marche simple, mais qui est la méthode la plus puissante pour la resolution des èquations de differences (différentielles).

On ajoute les équations $\overline{\psi}_n = p\overline{\psi}_{n+1} + q\overline{\psi}_{n-1}$ multipliées respectivement par $z^n, n = 1, 2, ..$ On obtient ainsi une èquation pour la fonction $\overline{\psi}^*(z) := \sum_{n=1}^{\infty} z^n \overline{\psi}_n$:

$$\overline{\psi}^*(z) = \frac{p\overline{\psi}_1}{\Phi(z) - 1}$$

où $\Phi(z) = Ez^{Z_1} = pz^{-1} + qz$

De lors,

$$\psi^*(z) = \frac{1}{1-z} - \frac{zp\overline{\psi}_1}{p + qz^2 - z} = \frac{p - qz - pz\overline{\psi}_1}{(1-z)(p-qz)} = \frac{p - qz - z(p-q)}{(1-z)(p-qz)} = \frac{p}{p-qz}$$

car le numerator s'annule en 1 ("méthode du noyau") et donc $p\overline{\psi}_1 = p - q$.

De lors, $\psi_n = (q/p)^n$.

2.7 Equations differentielles et récurrences linéaires

L'étude des marches aleatoires et des processus Markoviens ramène souvent à des équations differentielles ou de récurrence linéaires. Le cas des coefficients constants est assez simple, car toutes les solutions peuvent être construites à partir des solutions basiques exponentielles e^{rx}.

Comme le cas des équations differentielles à coefficients constants est très bien connu, on rappelle ici seulement le cas de récurrences linéaires.

2.7.1 L'équation de récurrence linéaire à coefficients constants

Les deux équations de récurrence linéaire de deuxième ordre ci-dessous

$$au_{n+2} + bu_{n+1} + cu_n = 0, \tag{8}$$

$$av_{n+2} + bv_{n+1} + cv_n = d_n, \tag{9}$$

sont appelées **homogène** et **nonhomogène** respectivement.

L'équation homogène

Si les coefficients a, b et c sont constants, on sait qu'ils existent des solutions de la forme $u_n = x^n$ pour tout n (fonctions exponentielles). Pour trouver x on remplace x^n en (8) et on trouve que x doit satisfaire **l'équation auxiliaire** :

$$ax^2 + bx + c = 0. \tag{10}$$

Soient x_1 et x_2 les deux racines de l'équation de deuxième degré (10). On en déduit que la solution *générale* de (8) est toujours de la forme

1. Si $x_1 \neq x_2$
$$u_n = Ax_1^n + Bx_2^n,$$
2. Si $x_1 = x_2$,
$$u_n = Ax_1^n + Bnx_1^n,$$
avec des constantes A et B.

Dans les deux cas A et B doivent être déterminées à partir des conditions supplémentaires sur la frontière.

L'équation nonhomogène

La résolution du problème nonhomogène (9) comporte quatre pas:

1. Trouver un basis pour l'espace vectoriel des solutions de l'équation auxiliaire homogène (8), et donc la solution générale u_n pour cette équation.

2. Déterminer une solution particulière de (9), par exemple en utilisant une expression "essai" $\tilde{v}_n$ qui a la même forme générale que le membre droit d_n,, mais des coefficients non-déterminés. Par exemple, si d_n est un polynôme d'ordre k, on essaie un polynôme général d'ordre k.

3. Néanmoins, si votre expression d'essai a des termes qui sont inclus dans l'espace vectoriel des solutions de l'équation homogène obtenue au pas 1 (et donc qui vont être annihilés par l'operateur des différences), il faut multiplier l'expression d'essai par $n, n^2, ...$ jusqu'à ce qu'il n'y a plus des termes inclus dans cet'espace .

4. Aprés la decision de la forme d'essai, on trouvent les valeurs des coefficients de $\tilde{v}_n$ à partir de (9), par la méthode des coefficients non-déterminés.

5. La solution générale de (9) est de la forme $v_n = \tilde{v}_n + u_n$. On trouve finalement les coeficients encore non déterminés en u_n, en utilisant les conditions sur la frontière pour v_n.

Exemple 2.2 *Obtenez les formules analytiques des suites décrites par les relations de récurrence ci-dessous, et vérifiez-les en utilisant les premiers termes de la suite t_2, t_3.*

1. $t_i = 2t_{i-1} + i - 1, \quad t_0 = 0$

2. $t_i = 2t_{i-1} + 5 \cdot 2^i, \quad t_0 = 0$

3. $t_i = 3t_{i-1} - 2t_{i-2} + 2, \ t_0 = 0, t_1 = 2$

4. $t_i = 2t_{i-1} - t_{i-2} + 2, \ t_0 = 0, t_1 = 2$

Solution:

1. C'est une équation nonhomogène, alors nous aurons:

$$t_i = \tilde{t}_i + A2^i, \ \ \tilde{t}_i = c_1 i + c_2 \ \text{ avec } \ c_1 i + c_2 = 2(c_1 i - c_1 + c_2) + i - 1$$

$$\begin{array}{lrcl} \text{et alors} & c_1 & = & 2c_1 + 1 \quad \text{et } \ c_1 = -1 \\ & c_2 & = & -2c_1 + 2c_2 - 1 \quad \text{et } \ c_2 = 2c_1 + 1 = -1 \\ & \tilde{t}_i & = & -i - 1 \quad \text{Finalement,} \\ & t_0 & = & = 0 = -1 + A \quad \text{et } \ A = 1 \\ & t_i & = & -i - 1 + 2^i \end{array}$$

2. C'est une équation nonhomogène, alors:

$$t_i = \tilde{t}_i + A2^i, \ \ \tilde{t}_i = ci2^i \ \text{ avec } \ ci2^i = 2(c(i-1)2^i/2) + 52^i$$

$$\begin{array}{lrcl} \text{et alors} & c & = & 5, \quad t_i = 5i2^i + A2^i \quad \text{et finalement,} \\ & t_0 & = & = 0 = A \quad \text{et } \ A = 0 \\ & t_i & = & 5i2^i \end{array}$$

3. C'est une équation de différences nonhomogène et l'équation quadratique attachée a les racines $1, 2$, alors nous aurons:

$$t_i = \tilde{t}_i + A_1 2^i + A_2, \;\; \tilde{t}_i = ci \;\text{ avec }\; ci = 3(ci - c) - 2(ci - 2c) + 2$$

$$\begin{array}{lrcl} \text{et alors} & c & = & -2 \quad \text{et } \; c_2 = 2c_1 + 1 = -1 \\ & \tilde{t}_i & = & -i - 1 \quad \text{Finalement,} \\ & t_0 & = & = 0 = -1 + A \quad \text{et } \; A = 1 \\ & t_i & = & -i - 1 + 2^i \end{array}$$

4. C'est une équation de différences nonhomogène dont les racines de l'équation quadratique attachée sont confondues égales à 1 donc nous aurons:

$$t_i = \tilde{t}_i + A_1 + A_2 i, \;\; \tilde{t}_i = c_1 i + c_2 \;\text{ avec }\; c_1 i + c_2 = 2(c_1 i - c_1 + c_2) + i - 1$$

$$\begin{array}{lrcl} \text{et alors} & c_1 & = & 2c_1 + 1 \quad \text{et } \; c_1 = -1 \\ & c_2 & = & -2c_1 + 2c_2 - 1 \quad \text{et } \; c_2 = 2c_1 + 1 = -1 \\ & \tilde{t}_i & = & -i - 1 \quad \text{Finalement,} \\ & t_0 & = & = 0 = -1 + A \quad \text{et } \; A = 1 \\ & t_i & = & -i - 1 + 2^i \end{array}$$

2.7.2 La méthode des fonctions génératrices

Exercice 2.7 *a) Calculez* $T(z) = \sum_{n \geq 0} T_n z^n$, *où*

$$T_0 = 0, \quad T_n = 2T_{n-1} + 1, n \geq 1$$

b) Trouvez T_n.

Exercice 2.8 *a) Calculez* $T(z) = \sum_{n \geq 0} T_n z^n$, *où*

$$T_0 = 1, \quad T_n = 2T_{n-1} + n - 1, n \geq 1$$

b) Trouvez T_n.

2.7.3 Equations differentielles et récurrences linéaires à coefficients polynomiaux

Dans le cas des coefficients polynomiaux la resolution est possible seulement dans des cas particuliers, et en utilisant toute une hierarchie des fonctions de base: hypergeometriques, d'Alembertiennes et Liouvilliennes.

3 Problèmes d'absorbtion/Dirichlet pour les chaînes de Markov

Sommaire: Nous étendrons ici l'approche de la section precedante pour les problèmes d'absorbtion de marches aleatoires, aux cas des chaînes de Markov de transition P. Brèvement, après avoir identifié l'operateur associé comme $G = P - I$, on a exactement les mêmes équations comme en avant, la seule difference étant que les solutions peuvent être rarement explicitées.

3.1 Les chaînes de Markov absorbantes

Définition 3.1 *Une chaîne s'apelle absorbante si tous ses états récurrents sont absorbants, i.e. $P_{i,j} = \delta_{i,j}$ pour chaque état i récurrent.*

Motivation: Parfois, une chaîne/marche est forcée de rester dans un sous-ensemble de son espace d'états initial par des diverses méchanismes de contrainte. Par exemple, une marche sur $\mathbb{Z}$ qui est contrainte à rester nonegative, donc en $\mathbb{N}$, pourrait être absorbée en 0 pour toujours, ou "réfléchie", i.e retournée en $\mathbb{N}$ dés qu'elle arrive dans le complement $\partial = \mathbb{Z} - \mathbb{N}$.

Le méchanisme de contrainte le plus simple est l'absorbtion. On appellera l'ensemble des états absorbants "cimetière" ∂; ceux-ci sont characterisés par des probabilités de transition $P_{i,j} = \delta_{i,j}$, $\forall i \in \partial, \forall j$.

3.2 Les problèmes de Dirichlet

Les problèmes de Dirichlet ont comme objet l'étude des temps de sortie N, de la distribution du point de sortie $X_N \in \partial$, et des diverses autres fonctions comme des prix finaux ou des coûts accumulés par la marche jusqu'au moment de son absorbtion en ∂.

Pour les marches aléatoires, on a vu que tous ces problèmes aboutissaient dans des équations de différences (ou différentielles, si l'espace d'états était $\mathbb{R}^d$). On verra maintenant que pour les chaînes de Markov, ces problèmes aboutissent dans des équations impliquant la matrice $G = P - I$. La méthode pour établir ces équations est toujours le conditionnement sur le premier pas.

Soit X_k une chaîne de Markov absorbante à espace d'états $S = \mathcal{T} \bigcup \partial = \{1, 2, ...I, C1, C2, ..\}$, où les états en $\mathcal{T}$ sont transitoires, et les états $\partial = \{C1, C2, ...\}$ sont absorbants. Soit

$\boldsymbol{Q}$	$\boldsymbol{P}^{(\mathcal{T},\partial)}$
$\mathbf{0}$	$\boldsymbol{I}$

la matrice de transition et soit $\boldsymbol{\beta}$ la distribution initiale.

Définition 3.2 *Soit X_t une chaîne absorbante, soit ∂ l'ensemble des états absorbants, et soit $\mathcal{T}$ le sous-ensemble (complémentaire) d'états transitoires. On appélera* **temps de sortie/absorbtion** *N le premier temps quand le processus X_t n'est plus en $\mathcal{T}$ (et donc est arrivé en ∂)*

$$N = \inf\{t : X_t \notin \mathcal{T}\} \in \{1, 2, ...\}$$

Remarque 3.1 *a) N est precisement le nombre de fois en $\mathcal{T}$, en incluant la position initiale.*

b) On pourrait èglement considérer le temps $\tilde{N} = N - 1 \in \{0, 1, 2, ...\}$ pass'e en les états transitoires, sans inclure la position initiale.

On s'intéressera dans la distribution et l'espérance des **variables d'absorbtion**, comme le temps d'absorbtion N, la position après absorbtion X_N et la position avant absorbtion X_{N-1}; on vera qu'elles sont calculables en utilisant des certaines relations entre leur distributions conditionnées par le point de départ.

Nous étudierons en suite plusieurs charactéristiques du temps N jusqu'au premier passage dans l'ensemble des états l'absorbants ∂:

1. Les espérances des temps d'absorbtion $\boldsymbol{n} = (n_i, i \in \mathcal{T})$, où $n_i = \mathbb{E}_i N$.
2. La distribution de N.
3. Les probabilités d'absorbtion dans les différents états absorbants (s'il y en a plusieurs).

3.3 Les espérances des temps d'absorbtion

Théorème 3.1 *a) Les espérances des temps d'absorbtion à partir des états transitoires $\boldsymbol{n}$ satisfont le* **système d'absorbtion**

$$\boldsymbol{n} = \boldsymbol{Q}\boldsymbol{n} + \mathbf{1}$$

b) Elles sont données explicitement par:

$$\boldsymbol{n} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}\, \mathbf{1}$$

c) Avec une distribution initiale $\boldsymbol{\beta}$, l'espérance $\bar{n} = \mathbb{E}_{\boldsymbol{\beta}} N$ du temps d'absorbtion est:

$$\bar{n} = \mathbb{E}N = \boldsymbol{\beta}(\boldsymbol{I} - \boldsymbol{Q})^{-1}\mathbf{1}$$

Corollaire 3.1 *Soit $\widetilde{\boldsymbol{G}} := \boldsymbol{Q} - \boldsymbol{I}$. Les espérances des temps d'absorbtion à partir des tous les états transitoires $\boldsymbol{n}$ satisfont le satisfont le système*

$$\begin{aligned} \widetilde{\boldsymbol{G}}\boldsymbol{n} + \mathbf{1} &= 0 \\ n_i &= 0, \quad \forall i \in \partial \end{aligned} \tag{11}$$

Ceci est notre premier exemple de système de Dirichlet. Rémarquez que la matrice $\widetilde{G}$ a seulement ses valeurs propres avec partie réelle negative, étant par conséquent inversible.

Demonstration: a) est équivalent au système

$$n_i = \sum_{j \in \mathcal{T}} Q_{i,j}(n_j + 1) + \sum_{j \notin \mathcal{T}} P_{i,j}^{(\mathcal{T},\partial)} * 1$$

qui généralise l'équation (1), et est obtenu par un conditionnement pareil.

b) est simplement la solution du système donné en a).

Exemple 3.1 *Soit une chaîne sur* $\{1,2\}$ *définie par la matrice de transition* $\left(\begin{array}{c|c} p & 1-p \\ \hline 0 & 1 \end{array}\right)$ *avec* $X_0 = 1$ *(i.e., la loi initiale est* $\boldsymbol{\mu}_0 = (1,0)$*).*

Soit N *le nombre des transitions jusqu'à l'absorbtion, en partant du temps* 0

a) Quelle est la valeur de N *si* $X_0 = X_1 ... X_{k-1} = 1$ *et* $X_k = 2$*? Quel est l'espace d'états de* N*?*

b) Trouvez l'espérance $n = \mathbb{E}N$ *du nombre des pas* N *jusqu'à l'absorbtion, en partant du premier état (i.e.* $X_0 = 1, \boldsymbol{\beta} = (1,0)$*).*

Rqs: 1) Le système donné au point a) est encore plus fondamental que la formule explicite donnée en b), car l'inversion des matrices est rarement la meilleure solution pour résoudre un système; aussi, il existe beaucoup de variations de ce problème où le système (un peu différent) sera facile à obtenir, par la même méthode de **conditionnement sur le premier pas**. Ce problème fournit une illustration du fait que conceptuellement et numériquement, les systèmes d'équations sont plus utiles que leurs solutions explicites!

2)(*) La formule explicite

$$\boldsymbol{n} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}\, \mathbf{1} = \big[\sum_{i=1}^{\infty} (\boldsymbol{Q})^i\big]\mathbf{1}$$

a aussi une interprétation probabiliste importante. Remarquons d'abord **la décomposition en indicateurs**

$$N = \sum_{k=0}^{\infty} I_k$$

où I_k est l'indicateur d'être dans la partie transitoire au temps k. Donc, $n_i = \sum_{k=0}^{\infty} \mathbb{E}_i I_k$. Remarquons aussi la décomposition en indicateurs $I_k = \sum_{j \in \mathcal{T}} I_{k,j}$, où $I_{k,j}$ est l'indicateur d'être en position $j \in \mathcal{T}$ au temps k. Ces décompositions nous ramènent finalement à

$$n_i = \sum_{k=0}^{\infty} \sum_{j \in \mathcal{T}} \mathbb{E}_i I_{k,j} = \sum_{k=0}^{\infty} \sum_{j \in \mathcal{T}} (\boldsymbol{Q})^k_{i,j}$$

3.4 Les probabilités d'absorbtion

Définition 3.3 *Soit* X_t *un processus, soit* E *un sous-ensemble arbitraire de l'espace d'états , et soit* ∂ *son complémentaire. On appelera* **processus absorbé en l'ensemble d'arrêt** ∂ *le processus* $\tilde{X}_t$ *obtenu à partir de* X_t *en modifiant tous les états en* ∂ *en sorte qu'ils soient absorbants.*

Dans le prochaine exemple, nous utiliserons plusieurs ensembles d'arrêt.

Exemple 3.2 *Pour une* **marche aléatoire** $X_t,\ t = 0,1,2,...$ **sur le graphe papillon** *ci-dessous, calculer:*

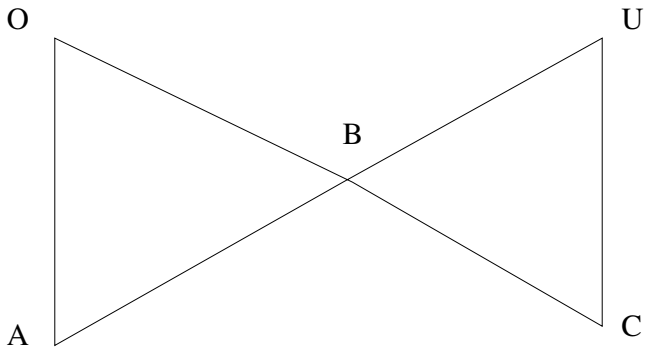


Figure 2: **Marche aléatoire simple sur le graphe papillon**

1. *L'espérance n_U en sortant de U du nombre de pas N jusqu'au noeud O.* Indication: *Utiliser la symmetrie.*

2. *Les probabilités stationnaires de chaque noeud.*

3. *L'espérance en sortant de O du nombre de pas $\tilde{N}_O$ jusqu'au premier retour à O.*

4. *La probabilité $p_A = \mathbb{P}_A\{X_N = O\} = \mathbb{P}_A\{N_O < N_U\}$, où $N = \min[N_U, N_O]$.*

Solution: 4) En résolvant le système d'absorbtion pour p_A, p_B, p_C, on trouve $p_A = 3/4, p_B = 1/2, p_C = 1/4$■.

Supposons qu'il y a plusieurs états absorbants à probabilités d'absorbtion $\boldsymbol{p}^{(j)}, j \in S-\partial$, qui donnent des "prix finals" $\boldsymbol{f} = \{f_j, j \in \partial\}$, posons $\hat{p}_i = \mathbb{E}_i f(N)$, et $\hat{\boldsymbol{p}} = \{\hat{p}_i, i \in S - \partial\}$ le vecteur de prix finals esperés. Le calcul de $\hat{\boldsymbol{p}}$ est le fameux **problème de Dirichlet**. Par exemple, pour $\boldsymbol{f}_j = \{0, ..., 0, 1, 0, ...\} = \{\delta_j(i), i = 1, 2, ...\}$ (avec le 1 sur la position j) on obtient les probabilités d'absorbtion $\hat{p}_i(j) = \mathbb{P}_i\{X_N = j\} = \mathbb{E}_i I_{\{X_N = j\}}$.

La théorie des chaînes pour les quelles tous les états récurrents sont absorbants peut être utilisée pour étudier n'importe quelle chaîne, en modifiant certaines transitions.

Théorème 3.2 *Le vecteur d'espérances du prix final $\boldsymbol{f}$ satisfait le* **système d'absorbtion**

$$\hat{\boldsymbol{p}} = \boldsymbol{Q}\hat{\boldsymbol{p}} + \boldsymbol{P}^{(\mathcal{T},\partial)}\boldsymbol{f}$$

En particulier

Théorème 3.3 *Les probabilités d'absorbtion $\hat{\boldsymbol{p}}^{(j)}$ dans un état absorbant fixe j satisfont le* **système d'absorbtion**

$$\hat{\boldsymbol{p}}^{(j)} = \boldsymbol{Q}\hat{\boldsymbol{p}}^{(j)} + \boldsymbol{p}^{(j)}_{(\mathcal{T},\partial)}$$

où $\boldsymbol{p}^{(j)}_{(\mathcal{T},\partial)}$ denote le vecteur des probabilités dans l'état absorbant j.

La matrice $\boldsymbol{P}^{(abs)}$ des probabilités d'absorbtion satisfait:

$$\boldsymbol{P}^{(abs)} = \boldsymbol{Q}\boldsymbol{P}^{(abs)} + \boldsymbol{P}^{(\mathcal{T},\partial)}$$

et donc

$$\boldsymbol{P}^{(abs)} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}\boldsymbol{P}^{(\mathcal{T},\partial)}$$

Corollaire 3.2 *Soit $\boldsymbol{G} := \boldsymbol{P} - \boldsymbol{I}$. Avec un prix final arbitraire $\boldsymbol{f}$, les prix finaux espérés $\boldsymbol{p}$ à partir de tous les états satisfont le* **système d'absorbtion**

$$\boldsymbol{G}\boldsymbol{p} = 0 \tag{12}$$
$$p_i = f_i, \quad \forall i \in \partial \tag{13}$$

Corollaire 3.3 *Avec une distribution initiale $\boldsymbol{\beta}$, avec un prix final arbitraire $\boldsymbol{f}$, on a l'espace d'états :*

$$\hat{\boldsymbol{p}} \;=\; \boldsymbol{\beta}(I - \boldsymbol{Q})^{-1})\boldsymbol{P}^{(\mathcal{T},\partial)}\boldsymbol{f}$$

Pour autres exemples des problèmes d'absorbtion, voir Ruegg, 2.6.4-2.6.5.

3.5 Les distributions des temps d'absorbtion (de type phase/"matrix geométric"

Les distributions de temps d'absorbtion, appelées aussi "distributions de phase", sont disponibles explicitement. Soit

$$\boldsymbol{p}(k) = (\boldsymbol{p}_{x,y}(k), x \in \mathcal{T}) := (\mathbb{P}_{x,y}\{N = k\}, x \in \mathcal{T}, y \in \partial\})$$
$$\boldsymbol{P}(k) = (P_{x,y}(k), x \in \mathcal{T}) := (\mathbb{P}_{x,y}\{N > k\}, x \in \mathcal{T}, y \in \partial\})$$

les matrices de dimension $|\mathcal{T} \times |\partial|$ ayant comme elements les probabilités jointes de k pas avant l'absorbtion, de départ en x et d'absorbtion en y.

Théorème 3.4 *a) Pour une chaîne de Markov absorbante à matrice de transition*

$\boldsymbol{Q}$	$\boldsymbol{P}^{(\mathcal{T},\partial)}$
$\boldsymbol{0}$	$\boldsymbol{I}$

les distributions du temps d'absorbtion N sont données par:

$$\boldsymbol{p}(k) \;=\; (\boldsymbol{Q})^{k-1}\boldsymbol{P}^{(\mathcal{T},\partial)}$$
$$\boldsymbol{P}(k) \;=\; (\boldsymbol{Q})^{k}\boldsymbol{1}$$

b) Avec distribution initiale $\boldsymbol{\beta}$, on a:

$$\mathbb{P}\{N = k\} \;=\; \boldsymbol{\beta}(\boldsymbol{Q})^{k-1}\boldsymbol{P}^{(\mathcal{T},\partial)}$$
$$\mathbb{P}\{N > k\} \;=\; \boldsymbol{\beta}(\boldsymbol{Q})^{k}\boldsymbol{1}$$

Exercice 3.1 *Démontrez le théorème. Ind: Les matrices $\boldsymbol{p}(k), \boldsymbol{P}(k)$ satisfont les recurrences:*

$$\boldsymbol{p}(k) = \boldsymbol{Q}\boldsymbol{p}(k-1), \qquad \textit{et } \boldsymbol{P}(k) = \boldsymbol{Q}\boldsymbol{P}(k-1)$$

qui ramènent (en itérant) au résultat:

Démonstration alternative: $I_{N=k} = \sum_{x_0,x_1,\ldots,x_{k-1}\in\mathcal{T},y\in\partial} I_{X_0=x_0,X_1=x_1,\ldots,X_{k-1}=x_{k-1},X_k=y\in\partial\in\mathcal{T}}$. Dès lors,

$$\mathbb{P}_{x_0}\{N=k, X_k=y\} = \sum_{x_1,\ldots,x_{j-1}\in\mathcal{T}} \boldsymbol{Q}_{x_0,x_1}\boldsymbol{Q}_{x_1,x_2}\ldots\boldsymbol{Q}_{x_{k-2},x_{k-1}}\boldsymbol{P}^{(tr)}_{x_{k-1},y} = (\boldsymbol{Q})^{k-1}\boldsymbol{P}^{(tr)}(x_0,y)$$

Les résultats sont obtenus en prenant somme en y et somme en x_0, ponderé par les poids $\boldsymbol{\beta}$. ■

Exercice 3.2 *Démontrez que*

$$\mathbb{E}N = \sum_k \mathbb{P}\{N>k\}$$

Rémarque: Ce théorème nous fournit une deuxième démonstration du Théorème 3.1 c):

$$\mathbb{E}N = \sum_k \mathbb{P}\{N>k\} = \boldsymbol{\beta}(\boldsymbol{I}-\boldsymbol{Q})^{-1}\mathbf{1}$$

Exercice 3.3 *Soit la matrice de transition* $\left(\begin{array}{cc|c} q_1 & 0 & p_1 \\ 0 & q_2 & p_2 \\ \hline 0 & 0 & 1 \end{array}\right)$ *et la distribution initiale* $(\beta_1,\beta_2,0)$. *Trouvez l'espérance et la distribution de* N.

Exercice 3.4 *Soit la matrice de transition* $\left(\begin{array}{cc|c} q_1 & p_1 & 0 \\ 0 & q_2 & p_2 \\ \hline 0 & 0 & 1 \end{array}\right)$

a) Trouvez l'espérance de N.

b) Généralisez au cas des matrices de ce type (série) de taille $K+1$.

c) Soit $N = N_1 + N_2$, *où* N_i *est le nombre de fois qu'on reste en* i. *Montrez que la distribution de* N_1 *conditionné par* N *est uniforme et calculez la distribution de* N, *si* $q_1 = q_2 = q$.

Remarque 3.2 *On peut aussi résoudre l'exercice en remarquant que* N_i *sont des variables géométriques, et donc* N *est hypergéométrique (une somme des géométriques).*

Voir aussi Ruegg, 2.6.3.

Conclusion: Les distributions de type phase demandent le calcul des puissances/exponentielles de matrices. Ces expressions sont très vite obtenues par logiciels comme Matlab, etc, et leur formule a toutes les apparences d'une expression analytique; mais, comme pour la plupart des matrices, les valeurs propres ne sont pas accessible analytiquement, leur calcul demande en effet une évaluation numérique.

3.6 L'opérateur associé à une chaîne de Markov

L'opérateur associé (ou générateur) à une chaîne de Markov est la "matrice génératrice" $G = P - I$.

Définition 3.4 *Une matrice G satisfaisant*

1. $g_{ij} \geq 0$ *si* $i \neq j$, $g_{ii} \leq 0$ *et*
2. $g_{i,i} + \sum_{j \neq i} g_{i,j} = 0$

sera appellée **matrice génératrice**.

Une matrice G satisfaisant 1. et $g_{i,i} + \sum_{j \neq i} g_{i,j} \leq 0$ sera appellée **sous-génératrice**.

Exercice 3.5 *Pour une matrice (sous)stochastique arbitraire, la matrice $G = P - I$ est une matrice (sous) génératrice.*

Remarque 3.3 *Il est facile de verifier que la matrice $G = P - I$ a les mêmes vecteurs propres comme P, et que ses valeurs propres sont translatées par -1.*

Les équations de Dirichlet –voir ci dessus– concernant les chaînes de Markov en temps discret peuvent être formulées egalement en termes de P ou de G, mais l'avantage de la dernière formulation et qu'elle generalise pour le temps continu.

3.7 Une classification des quelques problèmes concernant les processus de M

Note:1) Les problèmes que nous venons de discuter concernent une chaîne de Markov avec une ou deux états absorbants. Soit Q la matrice sous-stochastique des probabilités de transition parmi les états transitiores. En mettant $G = Q - I$, et $F(S) = \mathbb{R}^S$ l'espace des fonctions definies sur l'espace des états transitoires S, les équations trouvées ici pour les fonctions $p, f, t, c, w \in F(S)$ sont:

$$(Gp)_x = 0, b_B = 1, b_0 = 0$$
$$(Gf)_x = 0, f_B = B, f_0 = 0$$
$$(Gt)_x + 1 = 0, t_B = 0, t_0 = 0$$
$$(Gc)_x + x = 0, c_B = 0, c_0 = 0$$
$$(Gw)_x + (1 - a^{-1})w_x = 0, w_B = 1, w_0 = 1$$

font tous intervenir le même opérateur $G = P - I : F(S) -> F(S)$.

La remarque importante resortant de ces examples est qu'une grande partie des problèmes rencontrés peuvent être classifiés comme apartenant à un nb. assez petit de types de problèmes, qui sont associés aux formes specifiques des conditions frontière et des termes inhomogènes. Cela permet, une fois que le générateur d'un processus et le type de problème ont été reconnu, une identification méchanique des équations associées – voir Corollaires 3.1, 3.2, **??**.

Nous avons analysé cette correspondance dans le cadre le plus simple des chaînes de Markov finies, et pour les types de problèmes les plus faciles à dériver et résoudre: ceux de Dirichlet, corréspondant aux **problèmes d'absorbtion**. En suite, nous examinerons cette correspondance pour des modèles plus complexes en temps continu, et avec des espace d'états dénombrable et même continu.

3.8 Exercices

1. Soit une chaîne absorbante definie par la matrice de transition $\left(\begin{array}{cc|c} p_1 & p_2 & 1-p_1-p_2 \\ 0 & p & 1-p \\ \hline 0 & 0 & 1 \end{array}\right)$

   et la distribution initiale (α_1, α_2).

   Trouvez l'esperance du nombre des pas N jusq'à l'absorbtion.

2. Pour une **marche aléatoire** $X_t,\ t = 0, 1, 2, ...$ **sur le graphe cubique** ci-dessous, calculer:

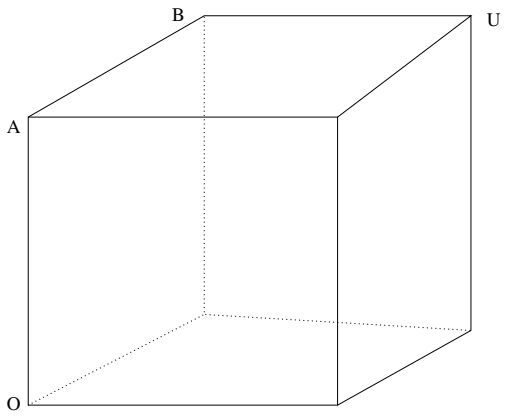


Figure 3: **Marche aléatoire simple**

   (a) L'ésperance en sortant de U du nombre de pas T_O jusq'au noeud O. *Indication:* Utiliser la symmetrie.

   (b) Les probabilités stationnaires du chaque noeud. **Indication:** Devinez la rèponse et montrez qu'elle satisfait le systême des équations d'équilibre.

   (c) L'ésperance en sortant de O du nombre de pas $\tilde{T}_O$ jusq'au premier retour à O.

   (d) La probabilité $p_A = \mathbb{P}_A\{X_T = U\}$, ou $T = \min[T_U, T_O]$.

   (e) La probabilité p_k en partant de O que la marche visite U exactement k fois $(k = 0, 1, 2, ...)$ avant le premier retour à O. Vérifier la somme $\sum_{k=0}^{\infty} p_k$.

3. **Rélations de récurrence**. Obtenez la formule analytique des suites décrites par les relations de récurrence ci-dessous, et verifiez-les en utilisant les premiers termes de la suite t_2, t_3.

   (a) $t_i = 2t_{i-1} + i - 1, \quad t_0 = 0$ (b) $t_i = 2t_{i-1} + 5 \cdot 2^i, \quad t_0 = 0$

   (c) $t_i = 3t_{i-1} - 2t_{i-2} + 2,\ t_0 = 0, t_1 = 2$ (d) $t_i = 2t_{i-1} - t_{i-2} + 2,\ t_0 = 0, t_1 = 2$

4. **La marche paresseuse:** Soit $X_n = X_0 + Z_1 + Z_2 + \cdots + Z_n$ une marche aléatoire sur $\mathbb{Z}$, avec (Z_n) une suite de variables aléatoires réelles indépendantes de même loi $P[Z_n = \pm 1]$ =p/q et $P[Z_n = 0]$ =1- p-q, avec $0 < p+q < 1$. Pour tout x de $\mathbb{N}$, on note par $\mathbb{E}_x$ l'espérance en commençant de x (conditionnant sur $X_0 = x$). Nous arrêtons le processus au temps d'arrêt T auquel le processus sort de l'intervalle $[0, K]$ pour $0 < K$ donnés.

   Obtenez l'équation de récurrence et les conditions frontière satisfaites par $p_x = \mathbb{P}_x\{X_T = K\}$, $f_x = \mathbb{E}_x X_T$, $t_x = \mathbb{E}_x T$, $c_x = E_x[\sum_0^T X(t)]$ et $w_x = \mathbb{E}_x a^T$.

   Résolvez les équations de récurrence qui en résultent pour a) p_x, b) f_x, c) t_x, d) c_x et e) w_x, quand $p = q < 1/2$.

5. a) Une mouche saute au hasard sur les sommets d'un triangle. Donnez une formule explicite, aussi efficace que possible, pour la probabilité qu'après n étapes, la mouche soit de retour au sommet dont elle est partie.

   b) Résolvez le même problème, au cas d'une marche cyclique avec probabilités de transition $b = P(1,2), c = P(1,3), a = P(1,1) = 1 - b - c$. En particulier, considérez le cas où la mouche a deux fois plus de chances à sauter dans le sens des aiguilles d'une montre.

   c) Généraliser au cas d'une marche cyclique sur les sommets d'un polygone avec k sommets, en utilisant votre language formel de choix.

6. Ruegg, 2.9:10,5,6,14,13,11.

7. Marc et un groupe de $n-1$ amis jouent un jeu. Chacun met un euro, et ensuite lance une monnaie biaisée, avec une probabilité de sortir "face" égale à p. La totalité de l'argent est partagé égalemment entre ceux qui ont obtenu face (s'il n'y a aucune, l'argent est donné à une oeuvre charitaire), et les piles perdent leur argent. a) Quelle est l'espérance du capital de Marc, après un tour? b) (*) Quelle est l'espérance du capital après un tour pour un joueur choisi aléatoirement ?

8. a) Quelle est la probabilité que la marche aléatoire simple est de retour an 0 après $2n$ pas? b) Approximer cette quantité par la formule de Stirling $\lim_{n\to\infty} \frac{n!}{(n/e)^n\sqrt{n}} = \sqrt{2\pi}$.

   c) (**) Démontrez la formule de Stirling.

3.9 Solutions

2) $X(t)$ représente une marche aléatoire simple sur le cube $[0,1]^3$, $\vec{0}$ est l'origine $(0,0,0)$, le coin oposé $(1,1,1)$ est noté par u. On remarque que dans toutes les questions les voisins de l'origine ont un role symetrique et cela nous permet de le noter par la même lettre $a = (0,0,1), \ldots$. De la même manière on appelle les voisins de u par $b = (0,1,1), \ldots$

a) Pour trouver $t_u = E_u[T_0]$, on résout le système:

$$\begin{aligned} t_u &= 1 + t_b \\ t_b &= 1 + \frac{2}{3}t_a + \frac{1}{3}t_u \\ t_a &= 1 + \frac{2}{3}t_b \end{aligned}$$

dont la solution est $t_a = 7, t_b = 9, t_u = 10$.

b) $E_0[\bar{T}_0]$ où $\bar{T}_0$ représente le temps esperé jusqu'à la prochaîne visite de 0 en commencant de 0, est donné par $1 + t_a = 1 + 7 = 8$. A remarquer que c'est exactement l'inverse de la probabilité "de long parcour" d'être à 0,ce qui est un résultat bien connu sur les temps de retours espérés.

c) $P_a[X(T) = u]$, s'obtient de la solution du système

$$\begin{aligned} p_a &= \frac{2}{3}p_b \\ p_b &= \frac{2}{3}p_a + \frac{1}{3} \end{aligned}$$

qui est $p_b = \frac{3}{5}, p_a = \frac{2}{5}$.

d) Soit p_k la probabilité d'avoir exactement k visites à $U = (1, 1, 1)$ avant de retourner à 0. Alors p_0 c'est le même que la probabilité commencant en A que la marche revient à 0 avant de visiter $(1, 1, 1)$, qui est $\frac{3}{5}$.

Pour $k = 1$ visite, "le chemin" observé seulement en O, l'état aprés B et U" est A,U,B,0. Donc, $p_1 = P_A[U, B, 0] = (\frac{2}{5})^2$, $p_2 = P_A[U, B, U, B, 0] = (\frac{2}{5})^2(\frac{3}{5})$, et en général $p_k = \frac{3}{5}p_{k-1} = (\frac{2}{5})^2(\frac{3}{5})^{k-1}, k \geq 2$. La distribution pour $k \geq 1$ est geometrique (verifiez que $\sum_{k=1}^{\infty} p_k = \frac{2}{5}$, comme il faut).

3) (a) C'est une équation nonhomogene, alors nous aurions:

$$t_i = \tilde{t}_i + A2^i, \;\; \tilde{t}_i = c_1 i + c_2 \;\text{ avec }\; c_1 i + c_2 = 2(c_1 i - c_1 + c_2) + i - 1$$

$$\begin{aligned} \text{et alors} \qquad c_1 &= 2c_1 + 1 \quad \text{et} \;\; c_1 = -1 \\ c_2 &= -2c_1 + 2c_2 - 1 \quad \text{et} \;\; c_2 = 2c_1 + 1 = -1 \\ \tilde{t}_i &= -i - 1 \quad \text{Finalement,} \\ t_0 &= = 0 = -1 + A \quad \text{et} \;\; A = 1 \\ t_i &= -i - 1 + 2^i \end{aligned}$$

(b) C'est une équation nonhomogène, alors:

$$t_i = \tilde{t}_i + A2^i, \;\; \tilde{t}_i = ci2^i \;\text{ avec }\; ci2^i = 2(c(i-1)2^i/2) + 52^i$$

$$\begin{aligned} \text{et alors} \qquad c &= 5, \quad t_i = 5i2^i + A2^i \quad \text{et finalement,} \\ t_0 &= = 0 = A \quad \text{et} \;\; A = 0 \\ t_i &= 5i2^i \end{aligned}$$

(c) C'est une équation differentielle nonhomogène et l'equation quadratique attachée a les racines $1, 2$, alors nous aurions:

$$t_i = \tilde{t}_i + A_1 2^i + A_2, \;\; \tilde{t}_i = ci \;\text{ avec }\; ci = 3(ci - c) - 2(ci - 2c) + 2$$

$$\begin{array}{rrcl} \text{et alors} & c & = & -2 \quad \text{et} \;\; c_2 = 2c_1 + 1 = -1 \\ & \tilde{t}_i & = & -i - 1 \quad \text{Finalement,} \\ & t_0 & = & = 0 = -1 + A \quad \text{et} \;\; A = 1 \\ & t_i & = & -i - 1 + 2^i \end{array}$$

(d) C'est une équation différentielle nonhomogène avec des racines confondues egales a 1 de l'equation quadratique attachée, alors nous aurions:

$$t_i = \tilde{t}_i + A_1 + A_2 i, \;\; \tilde{t}_i = c_1 i + c_2 \;\text{ avec }\; c_1 i + c_2 = 2(c_1 i - c_1 + c_2) + i - 1$$

$$\begin{array}{rrcl} \text{et alors} & c_1 & = & 2c_1 + 1 \quad \text{et} \;\; c_1 = -1 \\ & c_2 & = & -2c_1 + 2c_2 - 1 \quad \text{et} \;\; c_2 = 2c_1 + 1 = -1 \\ & \tilde{t}_i & = & -i - 1 \quad \text{Finalement,} \\ & t_0 & = & = 0 = -1 + A \quad \text{et} \;\; A = 1 \\ & t_i & = & -i - 1 + 2^i \end{array}$$

4) (a) Soit $(Gf)_x = p\ (f_{x+1} - f_x) + q\ (f_{x-1} - f_x)$ (formellement, la même expression comme dans le cas "non-paresseux", sauf que maintenant $p + q < 1$.

Les équations sont respectivement:

$$\begin{aligned} (Gp)_x = 0, p_K = 1, p_0 = 0 \\ (Gf)_x = 0, f_K = K, f_0 = 0 \\ (Gt)_x + 1 = 0, t_K = 0, t_0 = 0 \\ (Gc)_x + x = 0, c_K = 0, c_0 = 0 \\ (Gw)_x = (a^{-1} - 1)w_x, w_K = 1, w_0 = 1 \end{aligned}$$

Ces sont exactement les mêmes équations comme pour une marche non-paresseuse, sauf que l'opérateur est different.

(b) Pour p_x et f_x on obtient donc les mêmes équations comme pour une marche symmetrique avec $p = 1/2$, par exemple:

$$\begin{array}{rcl} 2pp_x & = & pp_{x+1} + pp_{x-1} \quad \text{for any } 1 \le x \le K - 1 \\ p_K & = & 1, \quad p_0 = 00 \end{array}$$

et donc les mêmes réponses $p_x = \frac{x}{K}, \quad f_x = x$ Pour $t_x = E_x[T]$ (temps de sortie esperé) on trouve:

$$\begin{aligned} t_x &= \frac{t_{x+1}}{2} + \frac{t_{x-1}}{2} + \frac{1}{2p} \quad \text{for any } 1 \le x \le K-1 \\ t_K &= 0, \quad t_0 = 0 \end{aligned}$$

avec solution $t_x = \frac{x(K-x)}{2p}$.

Remarque: Le fait que les équations sont identiques pour ces deux marches (paresseuse et non-paresseuse) n'est pas une coincidence. En fait, il s'avère que les réponses obtenues sont les mêmes por n'importe quel processus Markovien homogène, si on defini l'opérateur de la façon juste. Donc, la réponse pour tous les problèmes concernant espérances va impliquer un seul opérateur G (seulement les conditions frontière et la partie non-homogène changent d'un problème à l'autre)- en fait, la famille des processus aléatoires Markoviens est isomorphe à une famille d'opérateurs déterministes.

En plus, la structure des reponses en fonction de G est la même pour toutes les processus aléatoires Markoviens, malgré leur diversité; c'est un fait remarquable, qui démontre la puissance de ce concept.

5) a) L'équation de Chapman-Kolmogorov donne imédiatement une formule explicite: $P^n(1,1)$. On note aussi que les marches cycliques ont la matrice de transition P circulante, et donc une decomposition spectrale bien-connue explicite, qui utilise les racines (complexes) de l'unité. Mais, on peut faire mieux. La matrice P^n est aussi circulante, et contient donc seulement deux inconnues: $b_n = P^n(1,2), c_n = P^n(1,3)$. Soit $b = P(1,2), c = P(1,3), a = P(1,1) = 1-b-c$ les probabilités après un pas. On trouve la récurrence:

$$\begin{pmatrix} b_{n+1} - 1/3 \\ c_{n+1} - 1/3 \end{pmatrix} = \begin{pmatrix} a-b & c-b \\ b-c & a-c \end{pmatrix} \begin{pmatrix} b_n - 1/3 \\ c_n - 1/3 \end{pmatrix}$$

Le cas $b = c = 1/2$ et $a = b = c = 1/3$ donnent des récurences "decouplées". Le cas $b = 2/3, c = 1/3$ est plus difficile. En utilisant l'ordinateur, on rémarque que:

$$(b_n - 1/3, c_n - 1/3) = (1/3, 1/3) + 3^{-1-n/2} v_n$$

où $v_n = v_{n+12}$ est périodique.

1. a) X, le nombre total des faces a une distribution binomiale $B(n,p)$. Il y a deux possibiltés: - que Marc tire une pile et perd, dans quel cas son gain sera 0, et qu'il tire une face, dans quel cas le gain sera $Y = \frac{n}{1+X'}$ où X' a une distribution binomiale $B(n,p)$. Donc, l'espérance du gain est

$$\begin{aligned} Y = pE\frac{n}{1+X'} &= p\sum_{k=0}^{n-1} \frac{n}{1+k} C_{n-1}^k p^k q^{n-1-k} = p\sum_{k=0}^{n-1} \frac{n}{1+k} \frac{(n-1)!}{k!(n-k-1)!} p^k q^{n-1-k} \\ &= \sum_{k=0}^{n-1} \frac{n}{1+k} \frac{n!}{(k+1)!(n-k-1)!} p^{k+1} q^{n-1-k} = \sum_{j=1}^{n} C_n^j p^j q^{n-j} = 1 - q^n \end{aligned}$$

b) Le gain esperé d'un "joueur aléatoire" $Y = Y(X)$ est 0 si $X = 0$, a.p. q^n. Au cas contraire, le "joueur aléatoire" est gagnant avec probabilité $\frac{X}{n}$ et perdant avec probabilité $1 - \frac{X}{n}$. Le gain esperé est toujours $(1 - q^n)E[\frac{X}{n}\frac{n}{X}] = (1 - q^n)$.

Finalement, cet exercice suggère la question générale du calcul des "sommes binomiales", comme par exemple

$$S_n = \sum_{i=0}^{n-1} \frac{1}{(i+1)^2} C_{n-1}^i x^i = \sum_{i=0}^{n-1} \frac{1}{(i+1)^2} C_{n-1}^i x^i$$

où $x = \frac{p}{q}$. Parfois, ces sommes peuvent être déduites à partir de l'identité $(1 + x)^m = \sum_{i=0}^{m} C_m^i x^i$ en dérivant ou en intégrant. Mais, le proces d'integration n'abouti pas toujours à des sommes closes. Une somme $S_n = \sum_1^n f_n$ est une solution d'une relation de recurrence de premier ordre $S_n - S_{n-1} = f_n$ et donc la question de l'existence des formules closes pour f_n polynomes ou fonctions rationelles est un cas particulier de la question de l'existence des formules closes pour les recurrences avec coefficients polynomiaux.

Cette question est assez difficile, et le plus efficace est d'utiliser un logiciel symbolique. Ceux ci nous informent s'il y a des formules closes dans la famille relativement simple des solutions "d'Alembertiennes", ou si non.

3.10 Projet

Soit X_t une chaîne de Markov absorbante, soit ∂ l'ensemble de ses états absorbantes, soit B, A une decomposition de l'ensemble des états transitoires, et soit

$$\boldsymbol{p}(k, B) = (p_x(k, B), x \notin \partial)$$

où

$$p_x(k, B) := \mathbb{P}_x\{\text{exactement } k \text{ visites en } B \text{ avant l'absorbtion en } \partial\}, x \notin \partial$$

1. Quel type de distribution on trouve pour $p_x(k, B)$, quand $B = \{x\}$? (Specifiez les paramètres). Quel est le résultat pour la chaîne associé à:
$$\begin{pmatrix} 0 & 1-a & a & 0 & 0 \\ 1-b & 0 & b & 0 & 0 \\ x_1 & x_2 & 0 & x_4 & 1-x_1-x_2-x_4 \\ 0 & 0 & c & 0 & 1-c \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$ où $B = \{3\}$, et en particulier pour $a = b = c = 1/2, x_1 = x_2 = x_4 = 1/4$ ("le papillon").

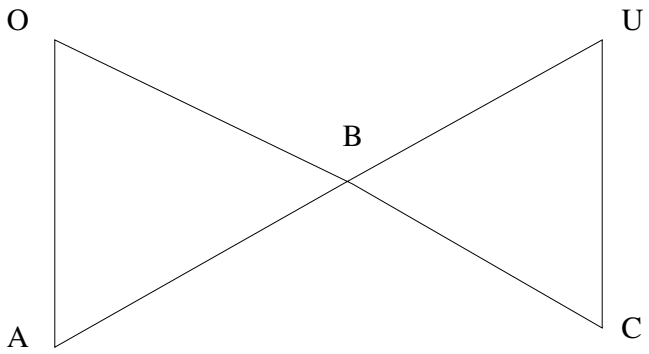


Figure 4: **Marche aléatoire simple sur le graphe papillon**

2. Pour B quelqonque, en conditionnant sur le premier pas, trouvez une relation entre les variables $p_x(k, B), x \in A, k \in \mathbb{N}$, et finalement une récurrence vectorielle $\boldsymbol{p}(k) = M\boldsymbol{p}(k-1)$, en spécifiant comment obtenir la matrice M à partir de la matrice P de transition de la chaîne. Vérifiez votre formule avec le cas $B = A$.

3. Retrouvez le résultat pour le "papillon généralisé" ci-dessus, dans le cas qu'on cherche la probabilité p_k en partant de $U = 5$ que la marche visite $O = 1$ exactement k fois ($k = 0, 1, 2, ...$) avant le premier retour à U) (les autres sommets seront libelés $A = 2, B = 3, C = 4$), **à partir de la formule générale**.

4. Considerez aussi le "papillon généralisé", en prenant $B = \{1, 2, 3\}$. Vérifiez pour cet exemple que la somme $\sum_{k=0}^{\infty} p_k(i, B) = 1, \forall i \in \{1, 2, 3\}$.

5. Ecrivez un program dans votre language de choix qui calcule $\boldsymbol{p}(k, B)$ et une approximation $\boldsymbol{p}(k, B) \approx c\lambda^k$ pour une chaîne et ensemble B arbitraires et démontrez sa performance sur les exemples 3.5, 3.6 (pages 23-24) et ensembles B de votre choix.

Solution:

1. Quand $|B| = 1$, on trouve une distribution geometrique $p_x(k, \{x\}) = \lambda^{k-1}(1-\lambda)$ où: $1 - \lambda = p_x(1, \{x\}) = Q(x, \partial(A) + \sum_{y \in A-B} p(x, y) P_y[T_{\partial(A)} < T_x] = Q(x, \partial(A) + \sum_{y \in A-B} p(x, y)(1 - P_y[T_{\partial(A)} > T_x])$, et $\lambda = Q_{B,B} + \sum_{y \in A-B} p(x, y) P_y[T_{\partial(A)} > T_x] = Q_{B,B} + Q_{B,A}(I - Q_A)^{-1} Q_{A,B}$, car pour $k \geq 2$ on a:

$$p_x(k, \{x\}) = \sum_{y \in A-B} p(x, y) P_y[T_{\partial(A)} > T_x] p_x(k-1, \{x\}) = \lambda p_x(k-1, \{x\})$$

Pour le papillon, $B = \{3\}$, $\lambda_B = x_1 + x_2 + cx_4 = 5/8$ et pour $B = \{1\}$, $\lambda_B = 3/8$.

2. Il est convenable de partager $\boldsymbol{p}_k = (\boldsymbol{a}_k, \boldsymbol{b}_k)$, où $\boldsymbol{b}_k = (p_x(k, B), x \in B)$, $\boldsymbol{a}_k = (p_x(k, B), x \in A, x \notin B)$. On peut supposer qu'il y a un seul état absorbant (en "collant ensemble" tous les états absorbants), et soit

$$P = \left(\begin{array}{cc|c} Q_A & Q_{A,B} & \boldsymbol{q}_A \\ Q_{B,A} & Q_B & \boldsymbol{q}_B \\ \hline 0 & 0 & 1 \end{array}\right)$$

la partition de la matrice de transition contenant les états dans l'ordre $A - B, B, \partial$. On a $\boldsymbol{b}_0 = 0, \boldsymbol{b}_1 = \boldsymbol{q}_B + Q_{B,A}\boldsymbol{a}_0$ et

$$\boldsymbol{a}_0 = \boldsymbol{q}_A + Q_A \boldsymbol{a}_0 \Longrightarrow \boldsymbol{a}_0 = (I - Q_A)^{-1} \boldsymbol{q}_A, \quad \boldsymbol{b}_1 = \boldsymbol{q}_B + Q_{B,A}(I - Q_A)^{-1} \boldsymbol{q}_A$$

Pour $k \geq 2, x \in B$, $p_x(k, B) = \sum_{y \in A} P(x, y) p_y(k-1, B)$, et donc

$$\boldsymbol{b}_k = Q_B \boldsymbol{b}_{k-1} + Q_{B,A} \boldsymbol{a}_{k-1}$$

tant que pour $x \notin B, k \geq 1$, $p_x(k, B) = \sum_{y \in A} P(x, y) p_y(k, B)$ et donc

$$\boldsymbol{a}_k = Q_{A,B} \boldsymbol{b}_k + Q_A \boldsymbol{a}_k \Longrightarrow \boldsymbol{a}_k = (I - Q_A)^{-1} Q_{A,B} \boldsymbol{b}_k$$

Comme

$$\boldsymbol{b}_1 = (I_B - Q_B)\mathbf{1}_B - Q_{B,A}\mathbf{1}_A + Q_{B,A}(I - Q_A)^{-1}((I_A - Q_A)\mathbf{1}_A - Q_{A,B}\mathbf{1}_B) = (I_B - M)\mathbf{1}_B)$$

on trouve

$$\boldsymbol{b}_k = (Q_B + Q_{B,A}(I - Q_A)^{-1} Q_{A,B}) \boldsymbol{b}_{k-1} \Longrightarrow \boldsymbol{b}_k = M^{k-1} \left((I_B - M)\mathbf{1}_B\right)$$

où $M = Q_B + Q_{B,A}(I - Q_A)^{-1} Q_{A,B}$ est la matrice de transition de la "chaîne induite" sur B (où "complement de Shur" de A en Q).

Quand $B = A$, on retrouve $\boldsymbol{b}_k = Q_B^{k-1} \boldsymbol{q}_B$.

3. En résolvant le système d'absorbtion pour p_A, p_B, p_C, on trouve $p_A = 3/4, p_B = 1/2, p_C = 1/4$. 5) Soit $p_{A,k} = \mathbb{P}_A\{\text{exactement } k \text{ visites en } U \text{ avant le retour en } O\}$, avec $p_{B,k}, p_{C,k}$ définies pareillement, et $\boldsymbol{p}_k = (p_{A,k}, p_{B,k}, p_{C,k})$.

   Ainsi, $\boldsymbol{p}_0 = (p_A, p_B, p_C)$ et $p_0 = \frac{1}{2}(p_{A,0} + p_{B,0}) = \frac{1}{2}(p_A + p_B)$.

   Pour $k \geq 1$, on trouve:

$$\begin{aligned}
\boldsymbol{p}_k &= \begin{pmatrix} 0 & 1/2 & 0 \\ 1/4 & 0 & 1/4 \\ 0 & 1/2 & 0 \end{pmatrix} \boldsymbol{p}_k + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/8 & 1/8 \\ 0 & 1/8 & 1/8 \end{pmatrix} \boldsymbol{p}_{k-1} \\
\iff \boldsymbol{p}_k &= \begin{pmatrix} 1 & -1/2 & 0 \\ -1/4 & 1 & -1/4 \\ 0 & -1/2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/8 & 1/8 \\ 0 & 1/8 & 1/8 \end{pmatrix} \boldsymbol{p}_{k-1} \\
\iff \boldsymbol{p}_k &= \begin{pmatrix} 0 & 1/8 & 1/8 \\ 0 & 1/4 & 1/4 \\ 0 & 3/8 & 3/8 \end{pmatrix} \boldsymbol{p}_{k-1}
\end{aligned}$$

   Les vecp à droite sont les colonnes de $\begin{pmatrix} 1 & 0 & 1/3 \\ 0 & -1 & 2/3/ \\ 0 & 1 & 1 \end{pmatrix}$, les valp correspondantes sont: $0, 0, 5/8$ et le vecp de PF à gauche est: $(0, 3/5, 3/5)$.

   Dés lors, $\boldsymbol{p}_k = (5/8)^{k-1}(p_B + p_C) \begin{pmatrix} 1/5 \\ 2/5 \\ 3/5 \end{pmatrix}$ et $p_k = (5/8)^{k-1}(p_B + p_C)3/10$

4 Le processus de Poisson

4.1 Rappels sur la distribution de Poisson et exponentielle

1. On considère deux variables aléatoires réelles X et Y indépendantes de loi de Poisson de paramètres $\lambda > 0$ et $\mu > 0$ respectivement.

   (a) Quelle est la loi de $Z = X + Y$? **Indication:** On peut utiliser la fonction génératrice des probabilitées $p_X(z) = \mathbb{E}z^X$, ou la fonction génératrice des moments $m_X(s) = \mathbb{E}e^{sX} = p_X(e^s)$.

   (b) Trouver la distribution d'une somme de n variables Poisson indépendantes de paramètre $\lambda_i, i = 1, ..., n$.

   (c) Pout tout entier $n \geq 0$, déterminer la loi conditionelle de X sachant que $Z = X + Y = n$.

   (d) Déterminer $E(X \mid X + Y)$.

(e) Quelle est la distribution de

$$S = \sum_{i=1}^{X} B_i$$

où B_i sont des va aléatoires Bernoulli de paramètre p?

2. **La loi exponentielle et la propriété de "manque de mémoire"**

Soit X une variable aléatoire réelle positive. Montrez que X suit une loi exponentielle si, et seulement si: $\forall t, h \geq 0$, $P([X - t \geq h] \mid [X \geq t]) = P[X \geq h]$. qu'on appelle la propriété de manque de mémoire.

Démonstration: Si X suit une loi exponentielle de paramètre $\lambda > 0$ alors pour tous $t, h \geq 0$ on a :

$$\begin{aligned} P([X \geq t+h] \mid [X \geq t]) &= \frac{P([X \geq t+h] \cap [X \geq t])}{P[X \geq t]} = \frac{P[X \geq t+h]}{P[X \geq t]} \quad (*) \\ P([X \geq t+h] \mid [X \geq t]) &= \frac{1 - F_X(t+h)}{1 - F_X(t)} \\ &= \frac{1 - \left(1 - e^{-\lambda(t+h)}\right)}{1 - (1 - e^{-\lambda t})} = e^{-\lambda h} = 1 - F_X(h) = P[X \geq h] \end{aligned}$$

Réciproquement , on suppose que $\forall t, h \geq 0$, $P([X \geq t+h] \mid [X \geq t]) = P[X \geq h]$ c'est-à-dire encore d'après $(*)$:

$$\forall t, h \geq 0 \ , \ P[X \geq t+h] = P[X \geq t] \, P[X \geq h]$$

d'où la fonction $\bar{F} = \bar{F}_X = 1 - F_X$ vérifie l'équation fonctionnelle

$$(**) \ \ \bar{F}(t+h) = \bar{F}(t) \, \bar{F}(h) \ \text{ pour tous } t, h \geq 0$$

En prenant logarithmes, on trouve que la fonction $f(x) = \log(\bar{F}(x)$ vérifie l'équation fonctionnelle

$$(**) \ \ f(t+h) = f(t) + f(h) \ \text{ pour tous } t, h \geq 0$$

On utilise maintenant le résultat:

Théorème 4.1 *Une fonction monotone satisfaisant l'équation fonctionnelle*

$$(**) \ \ f(t+h) = f(t) + f(h) \ \textit{pour tous } t, h \geq 0$$

doit être linéaire, i.e.

$$f(t) = tf(1)$$

Démonstration: : A partir de $(**)$, on obtient que :

$$\forall m,n\in\mathbb{N}\ ,\ f\left(\frac{m}{n}\right)=\left(f\left(\frac{1}{n}\right)\right)^{m}=\left(\left(f\left(1\right)\right)^{\frac{1}{n}}\right)^{m}=\left(f\left(1\right)\right)^{\frac{m}{n}}$$

Montrons que $f\left(1\right)\neq 0$: si $f\left(1\right)=0$ alors d'après ce qui précède f est nulle sur $\mathbb{Q}^{+}$, or on sait que pour tout réel $x>0$, il existe r dans $\mathbb{Q}^{+}$ tels que $r\leq x$, comme f est décroissante, on aura alors $0\leq f\left(x\right)\leq f\left(r\right)=0$ donc $f\left(x\right)=0$ et $f=0$, ce qui est faux.

Par conséquent les fonctions f et $x\mapsto\left(f\left(1\right)\right)^{x}=e^{x\ln f(1)}$ coïncident sur $\mathbb{Q}^{+}$, comme ces fonctions sont continues sur $\mathbb{R}^{+}$, on peut alors affirmer que $\forall x\geq 0\ ,\ f\left(x\right)=e^{x\ln f(1)}$

On sait que $\lim_{x\to+\infty}f\left(x\right)=1-\lim_{x\to+\infty}F_{X}\left(x\right)=0$ donc $\ln f\left(1\right)<0$ et on peut écrire que

$$\forall x\geq 0\ ,\ F_{X}\left(x\right)=1-e^{-\lambda x}\text{ avec }\lambda=-\ln f\left(1\right)>0$$

et on en déduit que la loi de X est une loi exponentielle.

4.2 Le processus de Poisson multidimensionel

Soit λ une constante. On appelle champs (où processus) de Poisson homogène d'intensité λ sur $\mathbb{R}^{d}$ un ensemble des v.a. $N(A)$ index'ees par les sous-ensembles $A\subset\mathbb{R}^{d}$ tq:

1. **indépendance:** si les sous-ensembles $A_1,...,A_k\subset\mathbb{R}^{d}$ sont disjoints, alors les variables $N(A_1),...,N(A_k)$ sont indépendantes
2. **stationarité:** la distribution des variables $N(A+t),t\in\mathbb{R}^{d}$ ne dépend pas de t
3. $N(A)$ est une variable Poisson d'intensité $\mu=\lambda\ m(A)$, où $m(A)$ est la mesure de Lebesgue.

Rq:

1. Un processus ayant les propriétés (1) et (2) s'appelle processus de Levy.
2. En fait, la troisième condition peut être remplacée par une condition beaucoup plus faible: $N(A)\in\mathbb{N}$, et la probabilité de plus d'un point dans des ensembles petits est négligeable, i.e. $P[N(A)\geq 2]=o(P[N(A)=1])$, quand $m(A)\to 0$.

Pour $t\in\mathbb{R}$, on arrive au processus de Poisson uni-dimensionel. Dans ce cas, on étudie les variables aleatoires indexées par les intervalles $A=[a,b]$, et on reserve le nom de processus de Poisson pour le processus $N(t),t\in\mathbb{R}$, défini par $N(t):=N([0,t])$. Les variables $N(A)=N(b)-N(a)$ sont appellées alors "incréments" du processus $N(t)$.

Définition 4.1 *Un processus qui a des acroissements independents (PAI) et pour le quel il existe une constante $\lambda>0$ tq pour tout $t>0$, la variable aléatoire $N_{t+s}-N_s$ suit une loi de Poisson de paramètre λt:*

$$p_k(t) := \mathbb{P}\{N_{t+s} - N_s = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

s'appelle processus de Poisson homogène.

Remarque 4.1 *Le nombre esperé d'arivées du processus de Poisson dans un interval de longueur t est:* $E[N(s+t) - N(s)] = \lambda t$.

Exemple 4.1 *Considèrons la distribution du nombre d'accidents de voiture dans une ville, par jour. a. Est ce-que ce nombre d'accidents pourra suivir approximativement une loi de Poisson? b. Trouver i. la probabilité qu'il y ait au plus trois accidents un jour donné. ii.le nombre moyen d'accidents (M) pour lequel la probabilité d'avoir trois accidents ou plus un jour donné est* < 0.05. *(On exprimera cette probabilité en fonction de l?inconnue M et on en déduira une équation qu'il faudra résoudre numériquement).*

Exercice 4.1 *Des clients arrivent dans une banque selon un processus de Poisson* $N(t), t \in \mathbb{R}$, *de paramètre* $\lambda = 1.2$ *(l'unité de temps est la minute).*

1. *Il arrive en moyenne combien de clients en 10 mn?*
2. *Donnez la loi de probabilité de N(2), le nombre de clients qui arrivent en deux minutes, et esquisser le graphique de* $p_k = P[N(2) = k]$, *pour k=0,1,2,3,4,5.*
3. *Donnez la probabilité que personne n'arrive durant 2 mn, et après ça que 3 personnes arrivent dans les 4 mn suivantes.*
4. *Donnez la probabilité* q_3 *que 3 personnes arrivent en 4 mn, vérifiez que* $q_3 = p_0 p_3 + p_1 p_2 + p_2 p_1 + p_3 p_0$, *et expliquez pourquoi.*

Exercice 4.2 *a) Montrez que la distribution du temps de la prochaine arrivée d'un processus de Poisson de paramètre* λ *est exponentielle a paramètre* λ.

b) Trouver la distribution du temps de la n-ième arrivée d'un processus de Poisson.

4.3 Processus de comptage et renouvellement en temps continu

Motivation: La plupart des phénomènes aléatoires demandent une modélisation et étude au cours du temps, comme les processus de comptage (où de naissances) $N_t, t \in \mathbb{R}_+$, tq les:

- appels arrivant dans un standard téléphonique
- arrivées de clients à un guichet
- survenue de pannes dans un parc de machines, ...

Définition 4.2 *Un processus* $(N_t)_{t \in \mathbb{R}_+}$ *est appelé processus de comptage s'il prend des valeurs dans* $\mathbb{N}$ *et s'il vérifie les trois propriétés suivantes :*

(i) $0 \leq s \leq t \Rightarrow N_s \leq N_t$
(ii) N_s *a des chemins "cadlag", i.e. continues à droite et avec des limites à gauche*
(iii) $P[N_s - N_{s_-} > 1] = 0$

Il s'agit donc des processus non décroissants et qui n'augmentent jamais avec plus d'une unité, et qui modélisent la réalisation d'une suite d'événements aléatoires d'un même type au cours du temps. Soit $a_i, i = 1, 2, ...$ les temps entre deux arrivés consecutives.

Définition 4.3 *Un processus de comptage pour le quel les temps entre deux arrivés consecutives sont des variables aleatoires i.i.d. s'appelle processus de renouvellement.*

Les temps de renouvellement (ou les temps de la n-iême arrivée) sont:

$$A_n = \sum_{i=1}^{n} a_i, n = 0, 1, 2, ...$$

Il est facile de voir que le nombre d'arrivées avant le temps t, i.e. le processus

$$(N_t)_{t\in\mathbb{R}_+} = \sup_k\{k : A_k \leq t\}$$

est un processus de comptage.

Théorème 4.2 *(admis) Le seul processus de renouvellement qui est aussi markovien est le processus de Poisson.*

4.4 Le processus de Poisson unidimensionel

On peut aussi définir le processus de Poisson homogène comme le processus de comptage associé à un processus de renouvellements, au cas où les temps a_i entre les arrivées sont exponentiels.

Définition 4.4 *Soit $a_i, i = 1, 2, ...$ une suite des v.a. i.i.d., à distribution exponentielle de paramètre λ ("interarrivées"), et soit $A_n = \sum_{i=1}^{n} a_i$ les temps de renouvellement. On appelera processus de Poisson homogène (en partant de $N_0 = 0$) le processus qui compte le nombre d'arrivées pendant $[0, t]$:*

$$N_{[0,t]} = N_t = \max\{n \in \mathbb{N} : A_n \leq t\}$$

Rq: La densité de A_n est la densité $\Gamma_{(n,\lambda)}(x)$.

Il existe plusieurs objets d'interêt associés à chaque processus de renouvellement:

1. l'âge courant $Z_1(t) = t - A_{N_t}$
2. le temps residuel courant $Z_2(t) = A_{N_t+1} - t$
3. le temps total courant $Z(t) = Z_1(t) + Z_2(t) = A_{N_t+1} - A_{N_t}$

Exercice 4.3 *Soit $N_\lambda(t)$ un processus de renouvellement de Poisson.*

a) Quelle sont les probabilités $P[Z_1(t) = t]$, et $P[Z_1(t) \in [s,t], \forall s \leq \infty$? Trouver la densité de $Z_1(t)$. Concluez que l'âge courant satisfait:

$$F_{Z_1}(x) = \begin{cases} 1 - e^{-\lambda} \text{ pour } x \leq t \text{ et} \\ F_{Z_1}(x) = 1 \text{ outrement} \end{cases}$$

b) Montrer que le temps residuel courant a une distribution exponentielle de paramètre λ.

c) Calculez l'espérance du temps total courant; et comparez la à l'espérance des temps entre les arrivées λ.

Les deux définitions (4.1), (4.4), sont équivalentes: voir les notes de Belisle. L'idée de la démonstration est presentée dans les exercices suivantes:

Exercice 4.4 *Soit $(N_t)_{t\geq 0}$ un processus défini par la Définition 4.4. Pour tout $t > 0$, la variable aléatoire N_t suit une loi de Poisson de paramètre λt, i.e.*

$$p_k(t) = \mathbb{P}\{N_t = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

Demo: a) Comme A_k est la somme de k variables exponentielles, elle a la densité Gamma $\Gamma(\lambda, k)$ donnée par: $f_{A_k}(t)dt = \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} \lambda dt$. On a aussi:

$$f_{A_k}(t)dt = \mathbb{P}\{A_{k-1} \leq t, A_k \geq t, A_k \leq t + dt\} = \mathbb{P}\{A_{k-1} \leq t, A_k \geq t\}\lambda dt = \mathbb{P}\{N_t = k-1\}\lambda dt$$

b) Calcul direct pour $k = 0, 1, ...$

Exercice 4.5 *Soit $(N_t)_{t\geq 0}$ un processus de Poisson unidimensionel d'intensité $\lambda > 0$, défini par la Définition 4.1. Montrez que les variables aléatoires suivantes: a_1, qui réprésente le temps jusqu'à la première arrivée, et a_2, qui réprésente le temps entre les deux premières arrivées, sont indépendantes et exponentielles de paramètre λ. Outrement dit, pour tout $t_1 > 0, t_2 > 0$,*

$$\mathbb{P}\{a_1 > t_1, a_2 > t_2\} = e^{-\lambda t_1} e^{-\lambda t_2}$$

Exercice 4.6 *La distribution exponentielle et l'independance des intervales a_i entre les arrivées implique la propriété de Markov. Plus specifiquement, montrez par exemple que:*

$$P[N_{t+s} \geq k+1 | N_t = k, N_{t-s_1} = k, \forall s, s_1 > 0]$$

Première méthode:

$$P[N_{t+s} \geq k+1 | N_t = k, N_{t-s_1} = k] = P[a_{k+1} \leq t+s-A_k | A_k \leq t-s_1, a_{k+1} > t-A_k]$$
$$= \frac{P[t-A_k < a_{k+1} \leq t+s-A_k, A_k \leq t-s_1]}{P[A_k \leq t-s_1, a_{k+1} > t-A_k]}$$
$$= \frac{\int_{x=0}^{t-s_1} f_{A_k}(x) P[t+s-x \geq a_{k+1} > t-x] dx}{\int_{x=0}^{t-s_1} f_{A_k}(x) P[a_{k+1} > t-x] dx}$$
$$= \frac{\int_{x=0}^{t-s_1} f_{A_k}(x)(e^{-\lambda(t-x)} - e^{-\lambda(t+s-x)} dx}{\int_{x=0}^{t-s_1} f_{A_k}(x) e^{-\lambda(t-x)} dx} = \frac{\int_{x=0}^{t-s_1} f_{A_k}(x) e^{-\lambda(t-x)}(1-e^{-\lambda s}) dx}{\int_{x=0}^{t-s_1} f_{A_k}(x) e^{-\lambda(t-x)} dx} = 1 - e^{-\lambda s}$$

Rq: Pour d'autres distributions, cette miraculeuse simplification n'est pas vraie [1].

Par consequent, le processus de Poisson est le seul processus de comptage Markovien.

Deuxiéme méthode (*). Pour démontrer des proprietées comme ci dessus, on utilise souvent la propriété "d'oubli de la durée de la dernière attente" de la loi exponentielle au temps t (voir Rappel **??**).

Ici on a:

$$P[a_{k+1} \leq t+s-A_k | A_k \leq t-s_1, a_{k+1} > t-A_k] = P[a_1 \leq s | A_0 = 0] = 1 - e^{-\lambda s}$$

4.5 Le processsus de Poisson comme limite des processus de Bernoulli

On peut aborder l'étude du processus de Poisson unidimensionel en discretisant le temps: on partagera chaque unité de temps en n unités "infinitesimales" de longueur $h = 1/n$, et en ignorant la possibilité des arrivés doubles. Ça remplace le processus d'arrivées en temps continu par un processus de Bernoulli (lancés d'une monnaie) en temps discret. En suite, en laissant $n \to \infty \Longleftrightarrow h \to 0$, on arrive au processus de Poisson. Le fait que le nombre (binomial) d'arrivées des succes dans un processus de Bernoulli en temps discret converge vers processus de Poisson en temps continu est une consequence de l'exercice suivant:

Exercice 4.7 *Considerez un processus d'arrives independentes dans des unités de temps de longueur $h = 1/n$, avec la probabilite d'une arrive per unite de temps egale a λ.*

1. *Montrez que la limite quand $n \to \infty$ d'une distribution binomiale $B(n,p)$ avec $p = \frac{\lambda}{n}$ est la distribution de Poisson avec espérance λ.*
2. *Soit $N_n^{(i)} i = 1, 2, ...$ les nombres d'intervalles separant les arrivées consecutives. Montrez que les temps $T_n^{(i)} = \frac{N_n^{(i)}}{n}$ entre les arrivées consecutives convergent vers des distributions exponentielles de paramètre λ, pour $i = 1, 2$.*
3. *Considerez un processus d'arrivées independentes, avec la probabilité d'une arrive per unite de temps (ou "taux") egale a λ. Trouver la distribution jointe des nombres d'arrives dans deux unitées de temps consecutives, en subdivisant chaque unitée en n periodes d'observation et en prenant $n \to \infty$.*

[1] La réponse depandra de t et de s_1 (on l'obtient en utilisant la formule $\bar{F}_{a_i}(x) = e^{-\int_0^x h(y)dy}$ où $h(y) = f_{a_i}(x)/\bar{F}_{a_i}(x)$ est le "taux de termination= hazard rate" de a_i).

4.6 Le processus de Poisson composé

Définition 4.5 *Le processus de Poisson composé*

$$S_t = \sum_{i=1}^{N_t} Z_i \tag{14}$$

est une somme des variables aléatoires i.i.d.(independentes et idéntiquement distribuées) Z_i, où N_t est un processus de Poisson, independant de Z_i.

Rqs: 1) Outrement dit, il s'agit des marches aléatoires avec increments qui arrivent après des temps exponentiels.

2) Si Z_i ont une distribution discrète $\boldsymbol{p} = (p_i, i \in \mathbb{Z})$, il s'agit simplement des marches aléatoires sur $\mathbb{Z}$, mais **en temps continu**; si en plus $p_i \neq 0$ ssi i est un voisin de l'origine, alors le processus (14) est une **marche aléatoire simple**.

In the next exercise we obtain formulas for the expectation, variance and cumulant generating fonctional of a compound Poisson process S_t (with bounded jumps) whose jumps' density is $f(x)$.

Exercise 4.1 a) Show that $\mathbb{E}S_1$, $\operatorname{Var} S_1$ for a compound Poisson process are given by:

$$\mathbb{E}S_1 = \lambda \mathbb{E}X_1$$
$$\operatorname{Var} S_1 = \lambda \mathbb{E}(X_1)^2$$

b) Compute the fonction génératrice des moments $\mathbb{E}e^{\theta S_t}$ compound Poisson process S_t (assuming bounded jumps) and show that $c(\theta)$ is given by:

$$c(\theta) = \lambda(M_{X_1}(\theta) - 1) = \lambda(\int_0^\infty e^{\theta = x} f(x)dx - 1) \tag{15}$$

Solution 4.1 a) Conditioning on the number of jumps we find

$$\mathbb{E}S_1 = \sum_{j=0}^{\infty} e^{-\lambda} \frac{(\lambda)^j}{j!} j\mathbb{E}X_1 = \mathbb{E}X_1 \mathbb{E}Po(1) = \lambda \int x f(x)dx$$

Similarly, the variance

$$\operatorname{Var} S_1 = \mathbb{E}(S_1)^2 - (\mathbb{E}S_1)^2 =$$
$$= \sum_{j=0}^{\infty} e^{-\lambda} \frac{(\lambda)^j}{j!} \mathbb{E}(\sum_{i=1}^{j} X_i)^2 - (\lambda \mathbb{E}X_1)^2$$
$$= \sum_{j=0}^{\infty} e^{-\lambda} \frac{(\lambda)^j}{j!} (j\mathbb{E}X_1^2 + j(j-1)(\mathbb{E}X_1)^2) - (\lambda \mathbb{E} = X_1)^2$$
$$= \lambda \mathbb{E}X_1^2 + \lambda(\mathbb{E}X_1)^2 - \lambda(\mathbb{E}X_1)^2 = \lambda \mathbb{E}X_1^2$$

b) Let $M_{X_1}(u) = \mathbb{E}e^{uX_1} = \int_0^\infty e^{ux} f(x)dx =$ denote the moment generating fonction of one jump. Conditioning on the = number of jumps we find

$\mathbb{E}e^{uS_t} = \mathbb{E}e^{u(\sum_{t=1}^{N_T} X_t)} = \sum_{j=0}^{\infty} = e^{-\lambda t} \frac{(\lambda t)^j}{j!} (M_{X_1}(u))^j = e^{\lambda t (M_{X_1}(u) - 1)} = e^{\lambda t (\int_0^\infty (e^{ux} - 1) f(x)dx)}$. Taking logarithms yields:

$$c(u) = \lambda(M_{X_1}(u) - 1)$$

Exercice 4.8 *Déterminer la fonction de covariance*

$$\operatorname{Cov}[X_t, X_{t+s}]$$

d'un processus de Poisson composé $X_t, t \geq 0$ (qu'on utilise pour modéliser le montant total des sinistres).

4.7 Conclusions

En conclusion, le processus de Poisson homogène unidimensionel de taux λ est un processus de comptage, où "l'avancé du compteur" se fait après des temps exponentielles de paramètre λ, et qui a plusieurs propriétés remaquables, comme:

- distribution exponentielle (donc "sans mémoire") de paramètre λ pour les intervales a_i entre les arrivées
- la propriété de Markov
- $(X_t)_{t\geq 0}$ est homogène, c'est-à-dire les increments $X_{[s,s+t]} = X_{s+t} - X_s$ ont une distribution indépendante du moment initial s:

$$\forall s, t > 0, \forall k \in \mathbb{N} \ , \ P\left[X_{s+t} - X_s = k\right] = P\left[X_t = k\right] \overset{not.}{=} p_k\left(t\right).$$

- $(X_t)_{t\geq 0}$ est un processus à accroissements indépendants (P.A.I. en abrégé), c'est-à-dire les v.a. $X_{I_i}, i = 1, ..., n$ sont indépendantes, si les intervalles $I_i, i = 1, ..., n$ sont disjoints. En particulier,

$$\forall s, t \geq 0 \text{ avec } s \leq t \ , \ X_t - X_s \text{ est indépendante de } X_u \text{ pour tout } u \leq s.$$

- $(X_t)_{t\geq 0}$ est un processus de Markov, et les probabilitées de transition satisfont $\mathbb{P}\{X_t = j | X_0 = i\} = \mathbb{P}\{X_t = j - i | X_0 = i\} = p_{j-i}(t)$.

Exemple d'application du processus de Poisson: La pêche! On note X_t le nombre de poissons pris par un pêcheur à la ligne dans l'intervalle de temps $[0, t]$. On fait les hypothèses suivantes :

($h1$) il y a un très grand nombre de poissons, de façon à ce qu'une prise n'influe pas sur la suite de la pêche

($h2$) l'appétit des poissons ne varie pas avec le temps.

Le processus $(X_t)_{t\geq 0}$ peut alors être considéré comme un processus de Poisson, car:

- $(X_t)_{t\geq 0}$ est bien un processus de comptage (il est assez clair qu'il est très peu probable de pêcher plusieurs poissons au même instant)
- $(X_t)_{t\geq 0}$ est homogène d'après ($h2$)
- $(X_t)_{t\geq 0}$ est un P.A.I. d'après ($h1$).

4.8 Exercices

1. a) Montrez que si les variables $\{a_i, i = 1, ..., n\}$ ont des distributions exponentielles independantes à paramètres $\lambda_i, i = 1, ..., n$, alors la variable $a = \min\{a_i, i = 1, ..., n\}$ a une distribution exponentielle de paramètre $\lambda = \sum_i \lambda_i$. b) Trouvez la distribution de la variable I qui donne l'index qui realise le minimum. c) Qu-est-ce qu'on obtient en superposant n processus de Poisson à paramètres $\lambda_i, i = 1, ..., n$?

2. Soit X_t un processus de Poisson de taux λ, les point du quel sont coloriés en K couleurs, independamment avec des probabilités $p_1, ..., p_K$, donnant naissance ainsi à K processus de comptage $X_1(t), ..., X_K(t)$. Montrez que $X_1(t), ..., X_K(t)$ sont des processus de Poisson independants, avec taux $\lambda p_1, ..., \lambda p_K$.

3. Des voitures passent sur une route selon un processus de Poisson de param. $\lambda = 1/mn$. a. Si 5% des voitures sont des Renault, montrer que le passage de ces voitures suit un processus de Poisson de param. 0.05. b. Sachant que 10 Renault sont passées en une heure, quelle est l'espérance du nombre total de voitures passées durant cette heure? c. Si 50 voitures sont passées en une heure, quelle est la probabilité que cinq d'entre elles soient des Renault?

4. **Le temps jusqu'à la première arrivée non-decouragée**. Soit N une variable géométrique, à distribution $\mathbb{P}\{N = k\} = (1-p)p^{k-1}, k = 1, ...,$, soit $T_i, i = 1, 2, ...$ des variables aleatoires i.i.d. exponentielles à paramètre μ, et soit $W = \sum_{i=1}^{N} T_i$ "le temps jusqu'à la première arrivée non-decouragée".

   (a) Trouvez la fonction génératrice des moments $m_{T_1}(s) = \mathbb{E}e^{sT_1}$ de T_1,

   (b) Trouvez la fonction génératrice des moments m_W de $W = \sum_{i=1}^{N} T_i$.

   (c) Quelle est la distribution de la variable aleatoire W?

5. **Ross, Exercice 63** Une société d'assurance paie pour les déclarations de sinistres dans le cadre des contrats d'assurance vie selon un processus de Poisson de taux $\lambda = 5$ par semaine. Si le montant payé pour chaque contrat suit une distribution exponentielle de moyenne 2000, quelles sont la moyenne et la variance des montants payés par la société d'assurances dans un intervalle de quatre semaines?

6. **Ross, Exercice 66** L'arrivée des clients à un distributeur de billets suit un processus de Poisson de taux 12 per hour. Le montant rétiré à chaque transaction est une variable aléatoire de moyenne 30 et écart-type 50 (un retrait négatif signifie que de l'argent a été déposé). La machine est utilisée 15 heures par jour. Approximez la probabilité que le montant total rétiré par jour soit infèrieur à 6000.

7. Belisle, TD4

Solutions:

1. c) Il s'agît d'un processus de comptage à interarrivées exponentielles, donc d'un processus de Poisson, a paramètre $\lambda = \sum_i \lambda_i$.

5 (a) $m_{T_1}(s) = \frac{\mu}{\mu - s} = \frac{1}{1 - s/\mu}$

   (b) $m_W(s) = \sum_{n=1}^{\infty}(1-p)p^{n-1}(\frac{\mu}{\mu-s})^n = (1-p)\frac{\mu}{\mu-s}\frac{1}{1-p\frac{\mu}{\mu-s}} = \frac{(1-p)\mu}{(1-p)\mu - s}$

   (c) W est donc exponentielle à paramètre $(1-p)\mu$, qui est précisement le taux des points acceptés.

4.9 La proprietè de Markov

Considérons d'abord, pour faciliter la notation, des processus X_t en temps continu, sur un espace d'états $\mathcal{E}$ fini ou dénombrable, appellés aussi **processus de sauts**. On note $\mathcal{E} = \{e_i; i \in I\}$ où $I \subset \mathbb{N}$, et on munit $\mathcal{E}$ de la tribu $\mathcal{P}(\mathcal{E})$, l'ensemble des parties de $\mathcal{E}$. Souvent, on identifie l'espace d'états $\mathcal{E}$ fini avec l'ensemble des indices I (i.e on appelle les états par leur indice).

Définition 4.6 ***-Proprietè de Markov***

Un processus $X = (X_t)_{t\geq 0}$ en temps continu et a éspace d'états $\mathcal{E}$ a la proprietè de Markov si, et seulement si ses probabilités conditionelles ne depend pas du passé que par le passé imediat, i.e.

$\forall 0 \leq t_0 < t_1 < \cdots < t_k < t \ , \ t_i \in \mathbb{R}, et \forall e_{i_0}, e_{i_1}, ..., e_{i_k}, e_i \in \mathcal{E}$

$$P([X_t = e_i] \mid [X_{t_0} = e_{i_0}, ..., X_{t_k} = e_{i_k}]) = P([X_t = e_i] \mid [X_{t_k} = e_{i_k}])$$

Un processus ayant la proprietè de Markov s'apelle processus de Markov.

Interprétation de la propriété de Markov: si on considère que le processus est indicé par le temps, cette propriété traduit le fait que le présent ne dépend du passé qu'à travers le passé immédiat.

Exemples: 2.3 et 2.1 de Ruegg, et les exemples de Belisle.

Définition 4.7 ***Matrice des transitions***

Pour tous $0 \leq s \leq t$, pour tous i, j dans I, et pour chaque processus de Markov, on définit les probabilités de transition par:

$$p_{ij}(s,t) = P([X_t = e_j] \mid [X_s = e_i]).$$

Définition 4.8 ***Homogeneité des transitions***

Un processus est dit homogène si, et seulement si :

$$\forall \ i, j \in I \ , \ \forall \ 0 \leq s \leq t \ , \ p_{ij}(s,t) = p_{ij}(0, t-s).$$

On note alors $p_{ij}(s,t) = p_{ij}(t-s)$, et la matrice $p_{ij}(t)$ est appellée matrice de transition après temps t.

Hypothèse de travail: $(H1)$ **On ne considérera ici que des processus homogènes.**

Un processus de Markov homogène est uniquement specifié par ses des probabilités de transition aprés temps t. Pour tout $t \geq 0$, on note par

$$P_t = (p_{ij}(t))_{i,j\in I}$$

la **matrice des probabilités de transition aprés temps** t.

L'exemple le plus simple des processus de Markov homogènes est fourni par les "chaînes de Markov" en temps discret et à espace d'états fini ou dénombrable. Le prochaîne exemple en ordre de complexité est celui de processus de sauts en temps continu.

En principe, toute la discussion précédente s'étend au processsus beaucoup plus compliqués, sur des espace d'états $\mathcal{E}$ arbitraires, en remplaçant comme d'habitude les événements "ponctuels" ci dessus par des événements mesurables arbitraires, ...

5 Les semigroupes (des matrices de transition) de Markov en temps continu

En temps discret, a cause de l'équation de Chapman-Kolmogorov, les matrice des probabilités de transition aprés temps n avait une structure de semigroup P^n, qui était engendré par la matrice P de transition aprés temps 1. On retrouve la même structure dans le cas continu:

Théorème 5.1 *Soit $(X_t)_{t\geq 0}$ une chaîne de Markov homogène à temps continu. Alors $(P_t)_{t\geq 0}$ est un semi-groupe stochastique, c'est-à-dire :*

(1) *$P_0 = I_d$*

(2) *$\forall s,t \geq 0$, $P_{t+s} = P_t P_s$ (équations de Chapman-Kolmogorov)*

(3) *$\forall t \geq 0$, P_t est une matrice stochastique (i.e. $\forall i,j \in I$, $p_{ij}(t) \geq 0$ et $\forall i \in I$, $\sum_{j\in I} p_{ij}(t) = 1$).*

Définition 5.1 *Une famille des operateurs satisfaisant les propriétés (1) et (2) du théorème 5.1 est appelée* **semi-groupe**. *Une famille satisfaisant aussi la propriété (3) est appelée* **semi-groupe stochastique**.

Démonstration : **(1)** Pour tous i,j de I , on a :

$$p_{ij}(0) = P([X_0 = e_j] \mid [X_0 = e_i]) = \delta_{ij} \text{ donc } P_0 = I_d.$$

(2) Pour tous i,j de I , pour tous $s,t \geq 0$, on a :

$$\begin{aligned} p_{ij}(t+s) &= P([X_{t+s} = e_j] \mid [X_0 = e_i]) \\ &= \sum_{k\in I} P([X_{t+s} = e_j] \cap [X_t = e_k] \mid [X_0 = e_i]) \end{aligned}$$

On rappelle que $P(A \cap B \mid C) = P(A \mid B \cap C)\, P(B \mid C)$ d'où :

$$\begin{aligned} p_{ij}(t+s) &= \sum_{k\in I} P([X_{t+s} = e_j] \mid [X_t = e_k] \cap [X_0 = e_i])\, P([X_t = e_k] \mid [X_0 = e_i]) \\ &= \sum_{k\in I} P([X_{t+s} = e_j] \mid [X_t = e_k])\, P([X_t = e_k] \mid [X_0 = e_i]) \\ &\quad \text{d'après la propriété de Markov} \\ &= \sum_{k\in I} P([X_s = e_j] \mid [X_0 = e_k])\, P([X_t = e_k] \mid [X_0 = e_i]) \\ &\quad \text{car } (X_t)_{t\geq 0} \text{ est homogène} \\ &= \sum_{k\in I} p_{ik}(t)\, p_{kj}(s) = (P_t P_s)_{ij} \quad \text{donc } P_{t+s} = P_t P_s. \end{aligned}$$

(3) Clair.

Pour charactériser les processus de Markov en temps continu, nous aimerions avoir une réprésentation des familles de matrices stochastiques de transition $P(t), \forall t \in \mathbb{R}_+$, qui satisfont l'équation de Chapman-Kolmogorov $P(t+s) = P(t)P(s), \forall t,s \in \mathbb{R}_+$

Dans le cas des semigroupes discrètes de matrices stochastique indicé par $t \in \mathbb{N}$, ça se faisait par la formule $P(t) = [P(1)]^n$.

Pour les processus de Markov en temps continu, il n'existe plus un "temps minimal" d'observation (comme le $t = 1$ du cas discrète); il est donc moins evident comment générer le semi-groupe des matrices de transition. Nous verrons que la réprésentation cherchée est

$$P(t) = e^{tG}$$

où la matrice G, appellée **générateur**, est liée á la matrice de transition infinitesimale P_h, avec $h \to 0$[2].

Exercice 5.1 *Calculer l'exponentielle e^{tG} d'un bloque de Jordan G.*

Nous supposera toujours la continuité du semigroupe en 0:

$$\lim_{t \to 0^+} P_t \to I_d \tag{16}$$

ou I_d denote l'identité, dans quel cas le semi-groupe $(P_t)_{t \geq 0}$ est dit standard.

Comme les propriétés (1) et (2) du théorème 5.1 sont celles d'une fonction exponentielle, on peut s'attendre par analogie avec le cas scalaire que $P_t = e^{tG}$ où le "générateur" G est la matrice donnée par:

$$G = \frac{dP_t}{dt}(0) = \lim_{t \to 0^+} \frac{P_t - I_d}{t} \tag{17}$$

(dérivée de $t \mapsto P_t$ en 0).

Il est donc naturel de tourner vers la matrice "des taux", définie par la rélation

$$P(t) \approx I - tG, \text{ pour } t \to 0_+ \quad \Longleftrightarrow \quad G := \lim_{t \to 0} \frac{P(t) - I}{t} = P'(0)$$

Définition 5.2 *On appelle générateur d'un semigroupe de Markov sur un espace d'états fini où denombrable la matrice $G = (g_{ij})_{i,i \in I} \overset{déf}{=} \lim_{t \to 0^+} \frac{P_t - I_d}{t}$.*

Plus précisement, la matrice G de "taux de transition" infinitesimales, ou "générateur", en partant d'un état initial arbitraire $X_0 = i$, est définie par:

$$\begin{cases} g_{ij} & = \lim_{t \to 0^+} \frac{P_{ij}(t)}{t} \quad \text{quand } i \neq j \\ -g_{ii} & = \lim_{t \to 0^+} \frac{1 - P_{ii}(t)}{t} \quad \text{quand } i = j \end{cases} \Longleftrightarrow \begin{cases} p_{ij}(t) = g_{ij} t + o(t)) \text{ quand } i \neq j \\ p_{ii}(t) = 1 + g_{ii} t + o(t) \end{cases}$$

[2]Comme definition de l'exponentielle d'une matrice, on peut utiliser celle basée sur la decomposition de Jordan, celle basée sur la serie Taylor, ou celle basée sur un systeme differentiel, i.e: $M_t = e^{tG}$ ssi

$$M_0 = I \text{ et } M'(t) = G\, M(t), \text{ ou } G = M'(0)$$

Ainsi, pour $j \neq i$ alors $g_{ij} = \lim_{t\to 0^+} \frac{p_{ij}(t)}{t} \geq 0$ est le **taux de passage** de l'état e_i à l'état e_j, et on a $g_{ij} \geq 0$ si $i \neq j$. Pour $j = i$, $-g_{ii} = \lim_{t\to 0^+} \frac{1 - p_{ii}(t)}{t} \geq 0$ est le **taux de sortie** de l'état e_i, pour $t \to 0$.

Note: En utilisant $P_t\mathbf{1} = \mathbf{1}$, on vérifie que $G\mathbf{1} = \mathbf{0}$, i.e. $g_{i,i} = -\sum_{j\neq i} g_{i,j}$.

Définition 5.3 *Une matrice G satisfaisant $g_{ij} \geq 0$ si $i \neq j$, $g_{ii} \leq 0$ et $g_{i,i} = -\sum_{j\neq i} g_{i,j}$ sera appellée* **matrice génératrice**.

Exemple 5.1 *Pour une matrice stochastique arbitraire, la matrice $G = P - I$ est une matrice génératrice.*

On verra que chaque matrice génératrice G défini un processus de Markov de sauts X_t (qui choisi son prochaine état $j \neq i$ à partir de $X_t = i$ avec probabilités proportionelles à $g_{i,j}, j \neq i$, aprés un temps exponentiel de paramètre $\lambda_i = \sum_{j\neq i} g_{i,j}$).

5.1 La propriété de Markov du processus de Poisson et son générateur

Soit $p_t(k) = P[N_\lambda(t) = k]$ la distribution du processus de Poisson[3] en començant de 0, et soit

$$P_t(i,j) = \begin{cases} p_t(j-i) & \text{si} j \geq i \\ 0 & \text{si} j < i \end{cases}, \quad i,j \in \mathbb{N}$$

sa matrice de transition.

Le fait que $p_t(k)$ satisfait:

$$p_{t+s}(n) = \sum_{k=0}^{n} p_s(k) p_t(n-k)$$

implique que les matrices de transition satisfont la propriété de semigroupe:

$$P_{t+s} = P_t P_s, \quad \forall s,t \in \mathbb{R}_+$$

Nous voyons ainsi que ce processus a la propriété de Markov en temps continu.

Le processus de Poisson est un des modèles les plus simples de processus ayant la **proprieté de Markov** en temps continu, et un des rares cas où les probabilités de transition ont des formules explicites. D'autre tels exemples sont fournis par les processus de saut sur un ensemble fini, qui nous aident à comprendre pourquoi les formules explicites pour les probabilités de transition sont si rares.

[3]Le processus de Poisson est a la fois le plus simple exemple d'un processus de sauts Markovien, est aussi un instrument théorique de démonstrations, grace a la possibilité de construire explicitement toutes les processus de sauts Markovien en partant des processus de Poisson. Nous seront plutôt interessé ici dans le premier aspect.

Exercice 5.2 *Montrer que pour un processus de Poisson et pour $h \to 0$ on a :*

$$p_k(h) = \begin{cases} \lambda h + o(h) \text{ si } k = 1 \\ o(h) \text{ si } k \geq 21 - \lambda h + o(h) \text{ si } k = 0 \end{cases}$$

Exemple 5.2 *La matrice génératrice du processus de Poisson est:*

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & \\ 0 & -\lambda & \lambda & 0 & \\ 0 & 0 & -\lambda & \lambda & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots \\ & \ddots & \ddots & & \ddots \end{pmatrix}$$

La façon la plus simple de definir de processus de Markov en temps continu est à travers leurs **taux de transition**, analogues au taux λ dans l'exercice ci-dessus. Parmi les plus étudiés sont les processus de naissances-mort, (analogues des marches aléatoires simples discrètes), qui n'augmente et ne diminue jamais avec plus d'une unité.

Un exemple proeminent est le nombre de clients à attendant leur service à un guichet, ou dans un centre d'appels. Les modèles Markoviens les plus étudiés sont le files d'attente M(λ)/M(μ)/s/s+K, où s répresente le nombre des serveurs, $K \in \mathbb{N}$ est le nb des places d'attente disponibles, et la lettre M (pour Markov) indique une distribution exponentielle de temps de service et entre les arrivées. Les taux de transition sont $g_{i.i+1} = \lambda_i, g_{i,i-1} = \mu_i = \min(s\mu, i\mu)$.

On s'interesse surtout dans le cas où le processus est contraint de rester nonegatif ou dans un interval, ce qui raménera à l'ètude de certaines équations de difference avec conditions de frontière. Par exemple, les distributions stationnaires des files d'attente sont assez faciles a calculer, mais les probabilités de transition sont (relativement) simple seulement dans les cas $s = \infty$ et $s = 1$.

Exemple 5.3 *La matrice génératrice de la file M/M/2 est:*

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & \\ \mu & -\lambda - \mu & \lambda & 0 & \\ 0 & 2\mu & -\lambda - 2\mu & \lambda & 0 \\ \vdots & 0 & 2\mu & -\lambda - 2\mu & \lambda \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

Exemple 5.4 *La matrice génératrice de la file M/M/3 est:*

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & \\ \mu & -\lambda - \mu & \lambda & 0 & \\ 0 & 2\mu & -\lambda - 2\mu & \lambda & 0 \\ \vdots & 0 & 3\mu & -\lambda - 3\mu & \lambda \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

5.2 Les équations différentielles de Chapman-Kolmogorov

Dans le cas des processus au temps continu, on a aussi deux formes alternatives **différentielles** de léquation de Chapman-Kolmogorov. différentielle de l'équation de semi-groupe Chapman-Kolmogorov:

Théorème 5.2 ***Equations de Kolmogorov****: Soit $(X_t)_{t\geq 0}$ un processus de Markov homogène à temps continu à valeurs dans $E = \{e_i \ ; \ i \in I\}$, de générateur G et de semi-groupe standard $(P_t)_{t\geq 0}$. Si E est fini ou si $\sup_{i\in I}(-g_{ii}) < +\infty$ alors on a:*

$$\begin{cases} P_t' = P_t\, G \text{ (équation directe)} \\ \\ P_t' = G\, P_t \text{ (équation inverse ou rétrograde)} \end{cases}$$

De plus, pour tout $t \geq 0$, $P_t = e^{tG}$.

Rémarque: Ces équations permettent de calculer, au moins numériquement, les matrices P_t.

Pour la démonstration, il suffit de remarquer qu'on a:

$$P_{t+h} \ - P_t = P_t\,(P_h \ - Id) = (P_h \ - Id)\,P_t$$

Le resultat est obtenu en divisant par h et en passant a la limite.

Notes: 1) Comme la matrice $P(t)$ doit etre stochastique pour chaque t, il découle que la partie réele des valeurs propres de G doit etre nonpositive.

2) Pour les processus à espace d'états infini, on parvient très rarement à obtenir des solutions analytiques pour leséquations de Kolmogorov.

Deux exceptions admettant résolution explicite sont les processus de Markov à deux états et le processus de Poisson.

5.2.1 Résolution des èquations Chapman-Kolmogorov pour le processus de Markov à deux états

On suppose que $(X_t)_{t\geq 0}$ est une chaîne de Markov homogène à valeurs dans

$E = \{e_1, e_2\}$, de générateur $G = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$ avec $\alpha, \beta \geq 0$.

On est bien dans les conditions d'application du théorème 1-2 (E est fini). La chaîne étant simple, on a deux méthodes pour la résolution des équations de Kolmogorov.

• <u>*Première méthode*</u> : on calcule e^{tG} pour tout $t \geq 0$. Pour cela, on va essayer de diagonaliser G. Le polynôme caractéristique P_G de G est donné par :

$P_G(\lambda) = \det(G - \lambda I_2) = \lambda(\alpha + \beta + \lambda)$.

→ <u>1^{er} cas</u> : $\alpha = \beta = 0$

On a $G = 0$ donc $e^{tG} = I_2$ pour tout $t \geq 0$.

→ <u>$2^{ième}$ cas</u> : $(\alpha, \beta) \neq (0, 0)$

La matrice G admet 2 valeurs propres distinctes donc est diagonalisable.

Les valeurs propres sont 0 (avec $\overrightarrow{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ comme vecteur propre associé) et $-(\alpha+\beta)$ (avec $\overrightarrow{v}_2 = \begin{pmatrix} \alpha \\ -\beta \end{pmatrix}$ pour vecteur propre associé).

En posant $P = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix}$, on a $P^{-1} = \dfrac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ 1 & -1 \end{pmatrix}$ et

$$G = PDP^{-1} \text{ où } D = \begin{pmatrix} 0 & 0 \\ 0 & -(\alpha+\beta) \end{pmatrix}.$$

On a alors :

$$P_t = e^{tG} = Pe^{tD}P^{-1} = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta+\alpha e^{-(\alpha+\beta)t} & \alpha-\alpha e^{-(\alpha+\beta)t} \\ \beta-\beta e^{-(\alpha+\beta)t} & \alpha+\beta e^{-(\alpha+\beta)t} \end{pmatrix} \text{ car } e^{tD} = \begin{pmatrix} 1 & 0 \\ 0 & e^{-(\alpha+\beta)t} \end{pmatrix}$$

(on peut remarquer que P_t est bien une matrice stochastique).

• <u>*Deuxième méthode*</u> (toujours pour $(\alpha,\beta) \neq (0,0)$) :

$$\text{on a : } P_t = \begin{pmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{pmatrix}$$

On a $P_t' = P_t G$ d'où : $\begin{cases} p_{11}'(t) = -\alpha p_{11}(t) + \beta p_{12}(t) & (1) \\ p_{12}'(t) = \alpha p_{11}(t) - \beta p_{12}(t) & (2) \\ p_{21}'(t) = -\alpha p_{21}(t) + \beta p_{22}(t) & (3) \\ p_{22}'(t) = \alpha p_{21}(t) - \beta p_{22}(t) & (4) \end{cases}$

De plus, comme P_t est une matrice stochastique, on a :

$$\begin{cases} p_{11}(t) + p_{12}(t) = 1 & (5) \\ p_{21}(t) + p_{22}(t) = 1 & (6) \end{cases}$$

Les équations (1) et (5) donnent : $p_{11}'(t) = -\alpha p_{11}(t) + \beta(1 - p_{11}(t))$ i.e.
$p_{11}'(t) + (\alpha+\beta)p_{11}(t) = \beta$ donc $p_{11}(t) = \dfrac{\beta}{\alpha+\beta} + Ce^{-(\alpha+\beta)t}$ avec $C \in \mathbb{R}$.

Or, on sait que $p_{11}(0) = 1$ donc $C = \dfrac{\alpha}{\alpha+\beta}$ d'où : $\boxed{p_{11}(t) = \dfrac{1}{\alpha+\beta}\left(\beta + \alpha e^{-(\alpha+\beta)t}\right)}$

Avec (5), on a alors : $\boxed{p_{12}(t) = \dfrac{1}{\alpha+\beta}\left(\alpha - \alpha e^{-(\alpha+\beta)t}\right)}$

En procédant de même, les équations (3) et (6) donnent :
$p_{21}'(t) = -\alpha p_{21}(t) + \beta(1 - p_{21}(t))$ i.e. $p_{21}'(t) + (\alpha+\beta)p_{21}(t) = \beta$
d'où $p_{21}(t) = \dfrac{\beta}{\alpha+\beta} + C'e^{-(\alpha+\beta)t}$ avec $C' \in \mathbb{R}$.

Comme $p_{21}(0) = 0$ on obtient $C' = -\dfrac{\beta}{\alpha+\beta}$ d'où : $\boxed{p_{21}(t) = \dfrac{1}{\alpha+\beta}\left(\beta - \beta e^{-(\alpha+\beta)t}\right)}$

Et avec (6), on a : $\boxed{p_{22}(t) = \dfrac{1}{\alpha+\beta}\left(\alpha + \beta e^{-(\alpha+\beta)t}\right)}$ et on retrouve P_t.

5.2.2 Résolution des èquations Chapman-Kolmogorov pour le processus de Poisson

Considerons le processus de Poisson Markov homogène à temps continu $(X_t)_{t\geq 0}$, de générateur

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & 0 \\ 0 & -\lambda & \lambda & 0 & \vdots \\ 0 & 0 & -\lambda & \ddots & 0 \\ \vdots & & 0 & \ddots & \ddots \\ \vdots & & \vdots & \ddots & \ddots \end{pmatrix}$$

On vera bientôt qu'il coincide avec le processus de Poisson d'intensité $\lambda \geq 0$ introduit ci-dessu.

Fixons l'état initial $i = 0$ et posons $P_{0,j}(t) = p_j(t)$. Les équations de Kolmogorov $P' = PG$ pour la première ligne sont

$$p'_j(t) = \sum_{k\neq j} p_k(t)\ g_{k,j} - p_j(t) \sum_{k\neq j} g_{j,k}$$

Note: En ecrivant ces équations comme $p_j(t+h) = \sum_{k\neq j} p_k(t)\ g_{k,j}\ h + p_j(t)(1 - h\ \sum_{k\neq j} g_{j,k})$ on voit que c'est encore des équations de conditionnement.

Pour le processus de Poisson, ces équations sont:

$$\begin{cases} p_0^{'}(t) = -\lambda p_0(t) \\ p_j^{'}(t) = \lambda\left(p_{j-1}(t) - p_j(t)\right) \text{ si } j \neq 0 \end{cases}$$

On sait aussi que $p_j(0) = 0$ si $j \neq 0$ et $p_0(0) = 1$.
On a donc un système différentiel pour déterminer les p_j, qui peut être resolue **reccursivement**:

$\rightarrow$ Pour $j = 0$:

$$\begin{cases} p_0^{'}(t) = -\lambda p_0(t) \\ p_0(0) = 1 \end{cases} \Rightarrow \underline{p_0(t) = e^{-\lambda t}}$$

$\rightarrow$ Pour $j = 1$:

$$\begin{cases} p_1^{'}(t) = \lambda\left(p_0(t) - p_1(t)\right) \\ p_1(0) = 0 \end{cases} \Leftrightarrow \begin{cases} p_1^{'}(t) + \lambda p_1(t) = \lambda e^{-\lambda t} \quad (*) \\ p_1(0) = 0 \end{cases}$$

$t \mapsto \lambda t e^{-\lambda t}$ est une solution particulière de $(*)$ et $t \mapsto Ce^{-\lambda t}$ (où $C \in \mathbb{R}$) est la solution générale de $p_1^{'}(t) + \lambda p_1(t) = 0$ donc $t \mapsto (C + \lambda t)\, e^{-\lambda t}$ est la solution générale de $(*)$.
De plus, on doit avoir $p_1(0) = 0$ donc $C = 0$ et on a $\underline{p_1(t) = e^{-\lambda t}\lambda t}$.

$\rightarrow$ Raisonnons alors par récurrence. Supposons que pour j dans $\mathbb{N}$, $p_j(t) = e^{-\lambda t}\frac{(\lambda t)^j}{j!}$ et déterminons p_{j+1}. On a :

$$\begin{cases} p_{j+1}^{'}(t) = \lambda\left(p_j(t) - p_{j+1}(t)\right) \\ p_{j+1}(0) = 0 \end{cases} \Leftrightarrow \begin{cases} p_{j+1}^{'}(t) + \lambda p_{j+1}(t) = e^{-\lambda t}\frac{t^j \lambda^{j+1}}{j!} \quad (**) \\ p_{j+1}(0) = 0 \end{cases}$$

$t \mapsto Ke^{-\lambda t}$ (où $K \in \mathbb{R}$) est la solution générale de $p'_{j+1}(t) + \lambda p_{j+1}(t) = 0$ et $t \mapsto \frac{\lambda^{j+1}t^{j+1}}{(j+1)!}e^{-\lambda t}$ est une solution particulière de $(**)$ donc $t \mapsto \left(K + \frac{\lambda^{j+1}t^{j+1}}{(j+1)!}\right)e^{-\lambda t}$ est la solution générale de $(**)$. Or $p_{j+1}(0) = 0$ d'où $K = 0$ et on obtient : $\underline{p_{j+1}(t) = \frac{(\lambda t)^{j+1}}{(j+1)!}e^{-\lambda t}}$ ce qui recupère la distribution Poisson (en eclaircissant la provenance de la terminologie "processsus de Poisson").

Nous venons de redémontrer que pour un processus de Poisson $(X_t)_{t\geq 0}$ d'intensité $\lambda > 0$, la variable aléatoire X_t suit pour tout $t > 0$ une loi de Poisson de paramètre λt.

Exercice 5.3 *Recalculez les probabilités de transition du processus de Poisson par la méthode des fonctions génératrices.*

Remarque: 1) Pour le cas des espaces d'états finies, les equations differentielles pour P_t calcule la fonction e^{tG}, et donc admettent toujours des solutions explicites en fonction de la decomposition Jordan de G. En plus, la fonction exponentielle de matrice est mise en oeuvre numeriquement dans toutes les logicielles modernes. Par contre, pour les espace d'états dénombrables, le calcul de P_t est assez difficile, même pour le cas le plus simple de la file M/M/1 [4].

2) La solution reccursive ci-dessus s'appuie sur la structure triangulaire du générateur, et peut-être aussi implementé pour tout processus de "Poisson composé de naissance pure" (permettant la naissance de jumeaux, etc).

Le calcul peut-être simplifié en començant par la substitution $p_k(t) = e^{-\lambda_k t}P_k(t)$, ou λ_k est le taux total de naissance en état k. Il s'avère que $P_j(t)$ sont des polynômes, faciles a obtenir reccursivement.

Exercice 5.4 *Recalculez les probabilités de transition du processus de Poisson par cette substitution.*

Solution: Les équations de Kolmogorov pour le processus de Poisson deviennent après la substitution:

$$\begin{cases} P'_0(t) = 0, P_0(0) = 1 \\ P'_j(t) = \lambda P_{j-1}(t), P_j(0) = 0 \text{ si } j \neq 0 \end{cases}$$

Il suit que $P_0(t) = 1$, et integrations succesives donnent $P_j(t) = \frac{(\lambda t)^j}{j!}$.

6 Semigroups of positive operators

To every Markov process $X(t)$ on some state space S, we may associate a family of "transition measures" $p_t(x, dy), x, y \in S$ (or kernels), where $t \in \mathbb{N}$ or $t \in \mathbb{R}_+$, depending on whether we have a discrete or continuous time Markov process. Then, we associate to this family of

[4] Dans ce cas, la transformé de Laplace de P_t est explicite, et on peut l'inverser, en arrivant a des combinaisons des fonctions Bessel.

measures a family of **positive "conditional expectation" operators** $P_t : \mathbb{R}^S -> \mathbb{R}^S$, defined on functions on S by

$$P_t(f)(x) = E_x f(X(t)) = \int_z p_t(x, dz) f(z) \tag{18}$$

and having the properties a) $f \geq 0$ implies $P_t f \geq 0$ and b) $P_t 1 \leq 1$.

The Markov property/"forgetting of the past" translates into the "semigroup" Chapman-Kolmogorov equation

$$p_{s+t}(x, dy) = \int_z p_s(x, dz) p_t(z, dy)$$

which may be written also as

$$P_t * P_s = P_{t+s},$$

where $*$ denotes the usual product of kernels.

In the simplest case of finite state spaces, semigroups may always be written as

$$P_t = \mathrm{e}^{tG} \iff P_t' = P_t G, P_0 = I \iff P_t' = G P_t, , P_0 = I \tag{19}$$

for some matrix G with nonnegative off-diagonal elements satisfying $G1 \leq 0$, called generator, and e^{tG} denotes the matrix exponential. Thus, the study of Markov processes/semigroups on finite state spaces (for example the computation of the matrix exponential e^{tG}) is an easy consequence of the classical spectral theory of matrices.

The equations governing the simple finite space case, like (19), have natural extensions to more complicated spaces, but their resolution and even interpretation may become considerably more challenging. Depending on the type of problem studied, one finds in the study of Markov processes the same main "heroes" as in other branches of analysis: the Cauchy/semigroup/Kolmogorov equation, harmonic and Poisson equations, Sturm-Liouville equations, etc...

The first challenge is pinpointing the generator, which may have several useful expressions. For example, in the case $S = \mathbb{R}^d$, and of "Levy processes" it turns out that the generator may be both written as an integro-differential operator and formally written as $G = \kappa(\nabla)$, where ∇ is the gradient operator and κ is the "cumulant exponent", i.e. logarithm of the moment generating function $\kappa(\theta) := \log\big(E_0\big[\mathrm{e}^{\theta X(1)}\big]\big)$; both representations may be used for solving equations.

6.1 Stationary/autonomous problems

The easiest equations to solve are the stationary ones, obtained formally by letting $t \to \infty$ in Kolmogorov's equation. The function to be determined may still appear "on the left or on the right" of the operator G, leading to either

$$G^* f(x) = 0 \quad \text{or} \quad G f(x) = 0. \tag{20}$$

The corresponding solutions, which may be viewed as left/right eigenfunctions associated with the maximal eigenvalue of G, are nonnegative by the Krein-Rutman theorem

(generalization of Perron-Frobenius), and have probabilistic interpretations of "invariant starting measure" and "harmonic function", respectively. For the second problem to admit nonconstant solutions, it is typically necessary to postulate that some part of the state space is absorbing, and there the values of the harmonic function are given, yielding thus a Dirichlet problem.

We will focus our exposition on a particular type of "eventual exit" problems, whose study originates in the classical gambler's and actuarial ruin problems. They lead to "autonomous" harmonic/Dirichlet equations:

$$G\psi(x) - q\psi(x) = 0, x \in S, \quad \psi(x) = g(x), x \in S^c \tag{21}$$

or to Poisson equations

$$Gf(x) - qf(x) + g(x) = 0, x \in S, \quad f(x) = 0, x \in S^c \tag{22}$$

In fact, often the simplest way to identify the generator of a Markov process is by formulating the equations for the exit probability $\psi(x)$ resulting from conditioning after one time step, obtaining thus the corresponding harmonic problem.

We will review in detail some classical examples: discrete random walks, Levy, birth and death processes (whose associated operators may be represented via tri-diagonal infinite matrices), and diffusion processes, and continue with exit problems for "spectrally-negative Wong-Pearson processes", which are the sum of a spectrally negative Levy process and of a diffusion with linear drift and quadratic variance rate. The "Wong-Pearson" pure diffusion case enjoys many interesting properties due to its connections with the hypergeometric special functions, and we will investigate to what extent this structure generalizes.

While the initial approach to exit problems for random walks was analytic, via Laplace transforms/generating functions, it became clear later that whenever analytic answers available, they may be derived more elegantly by using probabilistic concepts, like those of "skip-free/spectrally negative" processes and phase-type distributions – see for example Asmussen [**?**], Bertoin [**?**], Flajolet and Guillemin's [**?**], etc. Probability plays thus as usual the role of clarifying and unifying methods apparently different at first sight, for example by replacing analytic concepts like harmonic functions (which depend on the specific discrete/continuous/"q-nature" of the underlying space) by martingales.

6.2 Time dependent/transient problems

The ultimate challenge is the construction/approximation of the semigroup.

Exemple 6.1 *Consider one of the simplest continuous time Markovian models, the compound Poisson process, in which the arrival of independent jumps Y_i happens according to a Poisson process N_t (after independent exponential times) of rate λ:*

$$X(t) = \sum_{i=1}^{N_t} Z_i \tag{23}$$

Let L denote the convolution operator of adding one summand:

$$Lf(x) = E[f(x - Z_1) = \int_y f(x - z)P_Z(\mathrm{d}z)$$

where P_Z is the cdf of Z_i. By a Taylor expansion, this may also be expressed formally as a power series in the differentiation operator:

$$Lf(x) = E[f(x - Z_1) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!}(-1)^k E[Z_1^k] = p^*(-D)f(x) \tag{24}$$

where $p^*(\theta) := Ee^{\theta Z_1}$ and $D := \frac{\partial}{\partial x}$.

For the compound Poisson process X_t, conditioning on the number of jumps in $[0, t]$ yields now the semigroup representation

$$P_t f(x) = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{\lambda^k t^k}{k!} p^*(-D)f(x)$$

Finally, letting

$$\kappa(\theta) := t^{-1} \log[E_0 e^{\theta X_t}] = \lambda(Ee^{\theta Z_1} - 1)]$$

denote the **cumulant generating function/symbol/kernel** *of the compound Poisson process (23), we find that the generator is*

$$G = \kappa(D) = \lambda(p^*(-D) - 1).$$

The representation $G = \kappa(D)$ of the generator turns out to hold (in the sense of "pseudo-differential operators") for all Levy processes.

The compound Poisson is one of the three building blocks of a general Levy process, together with Brownian motion linear and linear drift, and the symbol of a Levy process combining this three components has the Levy-Khincin representation

$$G = \kappa(D) = \frac{\sigma^2}{2} D^2 + \mu D + \lambda(p^*(-D) - 1)$$

This suggests that the generators of all continuous time Markov may be represented as power series in D, i.e. $G = \kappa(x, D)$.

These representations turn out handy in deriving formally infinite expansions of the ruin probability in Neumann-series [?], [?].

In fact, all the usual spectral, asymptotic, and numeric approaches may be used, with the additional bonus that whatever the particular example of original interest was, the probabilistic interpretation guarantees its almost automatic extension to other Markov processes (on continuous, discrete, "q-" state spaces, function spaces, etc), simply by changing the generator!

In conclusion, the theory of general Markovian processes is a fascinating cocktail of analysis, algebra and probability. Especially intriguing is the interplay between analysis

and probability in the asymptotic regime, in several dimensions. Of course, exact solutions are very rarely available there, and even asymptotic expansions for problems with multiple boundaries, beyond the large deviations logarithmic approximation, are not well understood. To illustrate the gap between what we can and what we want, note that on one hand, this field is typically kept to two-dimensional logarithmic asymptotics by serious technical difficulties; on the other hand, there is the feeling that full asymptotic expansions might be obtainable by automatic methods – see the Pemantle-Wilson program. For example, note that in the absence of constraints, the large deviations rate of a Markov process coincides with the "symbol" of the operator, which is its formal expression as a power series in the partial derivative operators.

7 Les processus de naissance-et-de-mort

Les processus de Markov les plus simples sont les processus "de naissance et mort" qui se deplace sur $\mathbb{N}$ toujours aux voisins, "sans sauter position". Traditionellement, on denote les taux d'aller a droite par $\lambda_i, i \in \mathbb{N}$, et les taux d'aller a gauche par $\mu_i, i \in (\mathbb{N} - \{0\})$ (comme le processus est à valeurs dans $\mathbb{N}$, on a forcément $\mu_0 = 0$). On appelé les λ_i taux de naissance (ou de croissance) et les μ_i taux de mort (ou de décroissance).

Définition 7.1 *On appelle processus de naissance et mort une chaîne de Markov homogène à temps continu $(X_t)_{t\geq 0}$, à valeurs dans $\mathbb{N}$ telle que :*

$$\begin{aligned} \forall i,j \quad &\in \quad \mathbb{N} \ , \forall h > 0 \ , \\ P_{ij}(h) \quad &= \quad P([X_{t+h} = j] \mid [X_t = i]) = \begin{cases} \lambda_i h + o(h) \ \ si\ j = i+1^5 \\ \mu_i h + o(h) \ \ si\ j = i-1 \\ 1 - (\lambda_i + \mu_i) h + o(h) \ \ si\ j = i \\ o(h) \ \ sinon \end{cases} \end{aligned}$$

Autrement dit, le générateur d'un processus de naissance et mort de taux de naissance $\lambda_i, i \in \mathbb{N}$ et de taux de mort $\mu_i, i \in \mathbb{N}$ est donné par une **matrice tridiagonale***:*

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 \\ \vdots & 0 & & \ddots & \ddots \\ & \ddots & \ddots & & \ddots \end{pmatrix}$$

- Si tous les λ_i sont nuls, on parle de processus de mort.
- Si tous les μ_i sont nuls, on parle de processus de naissance.
  Exemple: processus de Poisson d'intensité λ ($\forall i \in \mathbb{N}$, $\lambda_i = \lambda$).

Proposition 1 *Un processus de naissance et mort de taux de naissance $\lambda_i, i \in \mathbb{N}$ et de taux de mort $\mu_i, i \in \mathbb{N}$ est irréductible si, et seulement si :*

$$\forall i \in \mathbb{N} \ , \ \lambda_i \neq 0 \ \textit{et} \ \forall i \in \mathbb{N}^* \ , \ \mu_i \neq 0.$$

Démonstration: si $\forall i \in \mathbb{N}$, $\lambda_i \neq 0$ et $\forall i \in \mathbb{N}^*$, $\mu_i \neq 0$ alors toutes les transitions sont possibles. Par contre, si l'un des réels λ_{i_0} ou μ_{i_0} est nul, une des transitions devient impossible, et on voit que la reunion des classes de communication contenant des elements a gauche de i_0 et celle contenant des elements a droite de i_0 ne peuvent pas communiquer. Alors, l'ensemble des classes de communication a au moins deux elements.
Remarque: • Deux domaines ou on utilise les processus de naissance et mort sont les **files d'attente** et les **modèles de population**, qui modélisent l'effectif X_t d'une population à l'instant t. Par exemple, pour modéliser une croissance/decroissane linéaires avec (ou sans) immigration et avec (ou sans) émigration, on utilisera:

$\rightarrow \forall i \in \mathbb{N}$, $\lambda_i = \lambda i + a$ avec $\lambda > 0$ et $a > 0$ (processus avec immigration) ou $a = 0$ (processus sans immigration)

$\rightarrow \forall i \in \mathbb{N}^*$, $\mu_i = \mu i + b$ avec $\mu > 0$ et $b > 0$ (processus avec émigration) ou $b = 0$ (processus sans émigration).

7.1 Les files d'attente Markoviennes

Le problème est le suivant : des "clients" arrivent en suivant une certaine loi de probabilité pour obtenir un "service" dont la durée est elle-même modélisée par une autre loi de probabilité. La file dispose d'un nombre $1 \leq s \leq \infty$ des "serveurs" et de K places supplementaires d'attente (ainsi,la capacité totale du système est $s + K$).

Dans cette application, la modélisation stochastique est plus aisée en temps continu. Soit $N_t \in \mathbb{N}, t \in \mathbb{R}$ le nombre total des clients dans le système (en attente ou en service) à l'instant t. Pour une file d'attente, les seuls transitions possibles de N_t sont l'ajout/départ d'un client.

Notation de Kendall. Kendall apelle un processus $N_t \in \mathbb{N}, t \in \mathbb{R}$ "file M/M/s/K" si N_t est un processus de naissance et mort, et

- il y a s serveurs
- la taille de la salle d'attente est de K places.
- les temps A_i entre les arrivée des clients et les temps de service B_i suivent des lois exponentielles (traditionellement, le taux d'arrivée/départ sont denoté par λ, μ) et sont independantes, le M étant une abbreviation de Markovien. Si les temps d'arrivée sont simplement independantes, on remplace le M par G(énérale), et s'ils sont de type phase, on remplace le M par Ph(ase).

La file d'attente M/M/1 a, en plus par rapport au processus de Poisson, aussi des "morts", le taux des quelles est traditionellement denoté par μ. Ainsi, elle est characterisé par les taux infinitesimales $g_{i,i+i} = \lambda, g_{i,i-i} = \mu, g_0 = \lambda, g_i = \lambda + \mu$ et sa matrice génératrice est:

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & \\ \mu & -(\lambda+\mu) & \lambda & 0 & \\ 0 & \mu & -(\lambda+\mu) & \lambda & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots \\ & \ddots & \ddots & & \ddots \end{pmatrix}$$

Autres exemples:

1) Pour la file M/M/1/K on a $\lambda_i = \lambda, i = 0, 1, ..., K, \lambda_{K+1} = 0$.

2) Pour la file M/M/s/K, on a $\mu_i = i\mu, i = 1, ..., s$, $\lambda_i = \lambda, i = 0, 1, ..., s+K-1, \lambda_{s+K} = 0$

Rémarque: Dans tous ces cas, le générateur est donné par une **matrice tridiagonale de Toeplitz**, i.e. avec diagonales constantes, au moins à une certaine distance des frontières.

Notes: 1) Comme les seuls processus en temps continu ayant la propriété de Markov doivent avoir des temps exponentielles entre les sauts, la seule file markovienne homogène est la file $M/M/s$.

2) Pour analyser un processus des sauts en temps continu, c'est naturel d'étudier aussi la **chaine discretisée** $\hat{X}_n$ obtenu en regardant seulement dans les moments des transitions (ce processus est parfois markovien, même quand le processus initial ne l'est pas). Pour la file M/M/1, la chaîne discretisée a une structure tres simple:

Exercice: Montrez que pour la file M/M/1:

a) $\mathbb{P}\{|TX_{n+1} = j|\tilde{X}_n = i\}$, quand $i \geq 1$ est donné par $\frac{\lambda}{\lambda+\mu}$ et $\frac{\mu}{\lambda+\mu}$ si $j = i \pm 1$ et par 0 outrement, et $\mathbb{P}\{\tilde{X}_{n+1} = j|\tilde{X}_n = 0\}$ est 1 si $j = 1$ et 0 outrement (plus tard nous allons généraliser cet exercice et donner une formule (**??**) pour la chaîne discretisé d'un processus des sauts en temps continu arbitraire.

b) Le temps d'attente entre les transitions est exponentiel à paramètre $\lambda + \mu$ conditionellement sur $\tilde{X}_n = i$ quand $i \geq 1$, est exponentiel à paramètre λ conditionellement sur $\tilde{X}_n = 0$.

En conclusion, la "chaîne discretisé" de la file M/M/1 a comme matrice de transition:

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & \\ \frac{\mu}{\mu+\lambda} & 0 & \frac{\lambda}{\mu+\lambda} & 0 & \cdots & \cdots \\ 0 & \frac{\mu}{\mu+\lambda} & 0 & \frac{\lambda}{\mu+\lambda} & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}$$

ce qui vérifie la formule (**??**).

7.2 Transformées de Laplace des probabilités transitoires de la file M/M/1

Pour la marche aléatoire simple en temps continu, la resolutions des èquations de Chapman-Kolmogorov est deja assez compliquée (impliquant des fonctions de Bessel). Par contre, on peut obtenir facilement les transformées de Laplace des probabilités transitoires. Nous illustrons maintenant les simplications apportées en appliquant la transformée de Laplace aux èquations de Chapman-Kolmogorov pour un processus de naissance mort de générateur

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & 0 & \vdots \\ 0 & \mu_2 & -\lambda_2 - \mu_2 & \ddots & 0 \\ \vdots & 0 & \mu_2 & \ddots & \ddots \\ \vdots & & \vdots & \ddots & \ddots \end{pmatrix}$$

, en partant de l'état initial $i = 0$. Posons $p_j(t) := P_{0,j}(t)$ et soit $p_j := p_j^*(s), j = 0, 1, 2, ...$ les transformées de Laplace.

Les équations de Kolmogorov $P' = PG$ pour la première ligne sont

$$\begin{aligned} p_0'(t) &= -\lambda_0\, p_0(t) + \mu_1\, p_1(t), \quad p_0(0) = 1 \\ p_j'(t) &= \lambda_{j-1}\, p_{j-1}(t) - (\lambda_j + \mu_j)\, p_j(t) + \mu_{j+1}\, p_{j+1}(t), \quad p_j(0) = 0, j = 1, 2... \end{aligned}$$

Pour les processus de naissance pure (ou mort pure), les équations ont une structure "triangulaire", qui permet de considerer et determiner d'abord p_0, en suite p_0, p_1, et et determiner p_1, et ainsi de suite (comme pour le processus de Poisson)...

En general par contre, nous sommes obligés de nous contenter d'obtenir les transformées de Laplace $p_j^*(s)$. Nous trouvons une recurrence homogène de deuxième ordre:

$$\begin{aligned} &\lambda_{j-1}\, p_{j-1}^* - (s + \lambda_j + \mu_j)\, p_j^* + \mu_{j+1}\, p_{j+1}^* = 0, \quad j = 1, 2... \\ &\mu_1 p_1^* - p_0^*\, (s + \lambda_0) + 1 = 0 \end{aligned}$$

La resolution des recurrences de deuxième ordre avec coefficients non-constants est bien sur toujours possible par iterations. Il existent aussi des formules analytiques utilisant les fractions continues.

L'inversion de la transformée de Laplace est possible par exemple par la formule de Post-Widder [6]

$$f(t) \sim (-1)^{k-1} \frac{s^k}{(k-1)!} \frac{d^{k-1}}{ds^{k-1}} \hat{f}(s)$$

où $s = k/t$.

Pour le cas de la file M/M/1 avec coefficients constants $\lambda_i = \lambda, \forall i \geq 0, \mu_i = \mu, \forall i \geq 1$, nous trouvons

$$p_n^*(s) = p_0^*(s)\alpha(s)^n$$

où $\alpha(s)$ est l'unique racine de $\mu\alpha^2 - (\mu + \lambda + s)\alpha + \lambda = 0$ qui est plus petite que 1 pour chaque $s \geq 0$, et $p_0^*(s) = (1 - \alpha(s)/s$ (car la normalisation implique $\sum_n p_n^*(s) = 1/s$).

Un autre cas abordable analytiquement est la file M/M/ı. Par contre, ce cas est plus facilement abordable en utilisant les propriétés de "coloriage" et "superposition" des variables de Poisson. Il resulte que les distributions des clientes sorties/en service dans la file M/G/∞ partant de 0 sont respectivement $Po(\lambda \int_0^t G(s)ds), Po(\lambda \int_0^t \bar{G}(s)ds)$ (exercice!).

7.3 Distribution stationnaire et comportement asymptotique

L'étude de files d'attente a comencé par le calcul d'Erlang (1921) de la probabilité de perte de la file M/M/s/0. En terminologie moderne, ça revienne au calcul de la probabilité stationnaire $\pi(s)$ de cette file.

La distribution stationnaire est une des charactéristiques les plus faciles à calculer des processus de Markov.

[6] Probabilistically, this amounts to replacing the fixed horizon t by an Erlang horizon with k stages and expectation t.

Définition 7.2 *On dit qu'un vecteur* $\pi = (\pi_i)_{i\in I}$ *est une* distribution stationnaire *si, et seulement si*

(1) π *est un vecteur de probabilité, c'est-à-dire :*

$$\forall i \in I \ , \ \pi_i \geq 0 \ et \ \sum_{i\in I} \pi_i = 1$$

(2) $\forall t \geq 0 \ , \ \pi = \pi P_t$.

Proposition 2 *Soit* $P_t = e^{tG}$ *un semigroupe de Markov. Alors, pour tout* $t \geq 0$*, il y a équivalence entre :*

(1) $\forall t \geq 0 \ , \ \pi = \ \pi P_t$ *(π vecteur propre de P_t a gauche pour la valeur propre 1)*

(2) $\pi G = 0$ *(π vecteur propre de G a gauche pour la valeur propre 0).*

Démonstration : $\pi G = 0 \Rightarrow \forall n \geq 1 \ , \ \pi G^n = 0$ donc $\forall t \geq 0 \ \pi \sum_{n\geq 1} \frac{t^n}{n!} G^n = 0$ d'où : $\pi \ e^{tG} = \pi P_t = \pi$

Réciproquement, supposons que $\forall t \geq 0 \ , \ \pi P_t = \pi$. On a alors : $\forall t \geq 0 \ , \ \pi P_t = \ \pi e^{tG} = \ \pi$. En dérivant par rapport à t , on obtient : $\pi G e^{tG} = 0$ et en prenant $t = 0$, on en déduit que $\pi G = 0$ d'où $\pi G = 0$.

Notes: Les équations de stationarité dans la forme

$$\sum_{j\neq i} \pi_j G_{j,i} = \pi_i g_i = \pi_i \sum_{j\neq i} G_{i,j} \tag{25}$$

ont l'interpretation d'une egalité entre le "flux entrant" en i et le "flux sortant" de i, où "flux" designe le taux de passage en/de i.

2) Parfois, un système plus simple "déquilibre partiel/local/ réversible"

$$\pi_i G_{i,j} = \pi_j G_{j,i}$$

est satisfait. Il est facile de verifier que ce système implique (25) et donc l'equilibre global $\pi G = 0$.

Par contre, le système déquilibre local est beaucoup plus simple à résoudre (quand il admet des solutions). Pour les processus de naissance et mort par exemple, le calcul de la distribution stationnaire se reduit toujours aux équations déquilibre local:

$$\pi_n \ \lambda_n = \pi_{n+1} \ \mu_{n+1}$$

Exercice 7.1 *Calculez la distribution stationnaire du nombre de clients dans les files suivantes.*

En particulier, obtenez la probabilité $\pi(s)$ *que les nouveaux arrivés ne doivent pas attendre (i.e. que le système a des serveurs libres), comme fonction de* $\rho = \frac{\lambda}{\mu}$*, ainsi que le nombre moyen de cients en attente.*

a) M/M/1 ($\pi_n = (1-\rho)\rho^n, \rho < 1$)
b) M/M/∞ ($\pi_n = e^{-\rho}\frac{\rho^n}{n!}$)
c) M/M/c/0 (la probabilité ($p(c)$ est la "fonction de perte" d'Erlang
c) M/M/c/K

Notes: 1) La file M/M/1 est "instable" (i.e. manque une distribution stationnaire de masse finie) si $\rho \geq 1$. La necessité de cette condition est evidente si on regarde la version chaîne en temps discret $X_n = \sum_1^n Z_i$, où $Z_i = \pm 1$ avec probabilités λ, μ tq $\lambda + \mu = 1$, et avec l'increment Z_i "censuré", i.e obligé d'être 0 au lieu de -1 quand $X_n = 0$.

Cette chaîne est stable ssi $EZ_i = \lambda - \mu < 0$. Pareillement, en temps continu, on vera que la file M/M/1 "libre" a se deplacer sur les negatifs $\tilde{X}_t$ satisfait $\mathbb{E}_x \tilde{X}_t = x + (\lambda - \mu)t$ et le manque d'ergodicité est assurée si $\lambda > \mu$, par la loi des grandes nb.

En généralisant, une marche $X_n = \sum Z_i$ en temps discret contrainte de rester en $\mathbb{N}$ sera stable ssi $EZ_i = \sum k p_k < 0$, et la condition analogue $\sum k\lambda_k < 0$ assure la stabilité en temps continu.

Exercice 7.2 *a) Formulez les équations d'équilibre pour la distribution stationnaire du nombre de clients en attente dans une file M(λ_1)/M(λ_{-1})/1, où il y aussi des arrivées des couples, avec taux λ_2.*
b) Donner la condition de stabilité.
c) (*) Résolvez les équations d'équilibre.

Théorème 7.1 *(admis) Soit $(X_t)_{t\geq 0}$ une chaîne de Markov à temps continu, homogène, de semi-groupe mélangeant* [7] *(ça implique la nonperiodicité et irréducibilité)$(P_t)_{t\geq 0}$.*
***(1)** Si π est une distribution stationnaire alors (le cas érgodique) on a:*
$\forall i, j \in I$, $\lim_{t\to+\infty} p_{ij}(t) = \pi_j$
***(2)** S'il n'existe pas de distribution stationnaire alors : $\forall i, j \in I$, $\lim_{t\to+\infty} p_{ij}(t) = 0$*

Remarques: 1) Pour un espace d'états fini, l'existence des solutions du système linéaire $\boldsymbol{\pi}\ G = \mathbf{0}$, ou $\boldsymbol{\pi}$ est un vecteur des probabilités, est toujours garantie par le théorème de Perron-Frobenius.

2) Le deuxième cas ci-dessus (qui n'aparaissait pas pour les espace d'états finis) arrive toujours si $EZ_i > 0$ en temps discret et $\mathbb{E}_0 \tilde{X}_t > 0$ en temps continu, par la loi des grands nombres. Ce cas implique une "fuite" du processus vers l'infini.

Reciproquement, pour pour un espace d'états infini, l'existance des solutions nonnulles (cas 1. ci-dessus) est assurée si le système "ne s'enfuit pas vers ∞".

[7] Un processus de Markov homogène en temps continu est dit mélangeant si, et seulement si :

$$\exists t > 0 \text{ tel que } p_{ij}(t) > 0, \forall i, j \in I$$

7.4 Mesures de performance des files d'attente

Quelques variables aléatoires interessantes pour les files d'attente qu'on veut étudier sont:

- La longueur Q_t de la file d'attente (des "clients" qui n'ont pas encore commençé leur service) à l'instant t et le nombre total N_t des clients dans le système (en attente ou en service) à l'instant t. Ces deux sont des processus de naissance et mort.
- La longueur W du temps d'attente dans la file (avant de commençér le service).
- La longueur totale T du temps dans le système (attente dans la file + service).

Pour ces quantités et aussi d'autres, on s'interesse en leur:

1. distribution stationaire, trouvé par "les équations d'équilibre" $\pi G = 0$.
2. les distribution "transientes" au temps t, en sachant l'état initial trouvé par "les équations de semigroupe", et
3. en la distribution cumulative de diverses "coûts" jusqu'aux temps d'arrêt T importants, comme celui de premier passage par 0, defini par $T_0 = \inf\{t \geq 0 : X_t \leq 0\}$. Par exemple, on s'interesse en le **"coût totale de vidage"** $c(x) = \mathbb{E}_x \int_0^{T_0} X_s ds$, et en le **coût de vidage discompté** $c_\delta(x) = \mathbb{E}_x \int_0^T e^{-\delta s} X_s ds$. Ces quantités sont trouvé par "les équations de Dirichlet/Feynman-Kac".

Comme le deuxième objectif est le plus difficile, permettant principalement seulement des approches numériques, ces sont le premier et troisième types des problêmes qui constituent l'objet d'étude principal de la théorie des files d'attente.

Exercice 7.3 *Calculez la fonction de distribution complementaire $\bar{F}(x)$ (probabilité que la variable soit plus grande que x) stationnaire du temps total de service et du temps d'attente, ainsi que le delai moyen dans les files:*
a) M/M/1 $(\bar{F}(x) = e^{\lambda-\mu)x}, \bar{F}_q(x) = \rho e^{\lambda-\mu)x}, \bar{W} = \frac{1}{\mu-\lambda}, \bar{W} = \frac{\rho}{\mu-\lambda}$
b) M/M/c $(\bar{F}(x) = \frac{1}{1-\rho/c} q_c e^{\lambda-c\mu)x}, \bar{F}_q(x) = \frac{1}{1-\rho/c} q_c e^{\lambda-c\mu)x}$

On peut verifier dans cette exemple les egalités $\bar{N} = \lambda \bar{W}, \bar{N}_q = \lambda \bar{W}_q$, un cas particulier de la loi de Little.

On peut aussi obtenir une formule pour le "travail stationnaire qui reste" $W(t)$ dans la file M/G/1 ($W(t)$ est un processus de Levy réflechi). On obtient la fameuse formule de Benes-Pollaczek-Khinchin, que nous verrons d'abord dans le contexte de l'actuariat.

7.5 La loi distributionelle de Little pour la file M/G/1 (*)

Exemple: la file M/M/1.

1. Trouvez la distribution stationnaire du nombre total N^* des clients dans la file M/M/1 et sa fonction génératrice.

2. Trouvez la distribution stationnaire du nombre Q^* des clients dans la file d'attente et sa fonction génératrice.

3. Montrez, en utilisant les fonctions génératrices, que

$$N^* = A(T) \quad \text{et} \quad Q^* = A(W)$$

ou $A(t)$ denote le nombre des clients arrivées dans un interval de longueur t. *Indication:* Posons $K(t,z) = \sum_{n=0}^{\infty} z^n \mathbb{P}\{A(t) = n\}$. Montrez d'abord que les rélations correspondantes entre les fonction génératrices $G_N(z), G_Q(z)$ de N, Q et les distributions de T, W, i.e.

$$G_N(z) = \int_0^\infty K(t,z) dF_T(t), \quad G_Q(z) = \int_0^\infty K(t,z) dF_W(t) \tag{26}$$

simplifient dans le cas d'arrivées Poisson avec $K(t,z) = e^{-\lambda t(1-z)}$ au

$$G_N(z) = \hat{f}_T[\lambda - \lambda z], \quad G_Q(z) = \hat{f}_W[\lambda - \lambda z] \tag{27}$$

ou $\hat{f}_T[s], \hat{f}_W[s]$ sont les transformées de Laplace de T, W.

4. Recalculez les fonction génératrices et transformées de Laplace dans les quatres premières questions, en partant des dernières deux lois (28) (appellées **les lois de Little**) et des rélations: $T = W + B, Q = N - 1_{\{N \geq 1\}}$ [8].

Solution: En transformant les deux dernières rélations, on arrive au système:

$$\begin{aligned} G_N(z) &= \hat{f}_T[\lambda - \lambda z], & G_Q(z) &= \hat{f}_W[\lambda - \lambda z] \\ \hat{f}_T[s] &= \hat{f}_W[s]\frac{\mu}{\mu + s}, & G_Q(z) &= \frac{G_N(z) - G_N(0)}{z} \end{aligned}$$

Alors:

$$G_N(z) = \hat{f}_W[\lambda - \lambda z]\frac{\mu}{\mu + \lambda - \lambda z} = G_Q(z)\frac{\mu}{\mu + \lambda - \lambda z} = (G_N(0) + \frac{G_N(z) - G_N(0)}{z})\frac{\mu}{\mu + \lambda - \lambda z}$$

En conclusion,

$$G_N(z)(z(\mu + \lambda - \lambda z) - \mu) = G_N(z)(z-1)(\mu - \lambda z) = G_N(0)(\mu(z-1)))$$

et $G_N(z) = \frac{G_N(0)}{1 - \rho z}$, d'ou (en faisant $z = 1$) $G_N(0) = 1 - \rho$, et N est geométrique $Geo(\rho)$. En suite, T est $Expo(\mu - \lambda)$ et W est $1 - \rho + \rho Expo(\mu - \lambda)$.

[8] Traditionellement, on appelait lois de Little les identités pour les ésperances: $\mathbb{E}Q = \lambda \mathbb{E}W, \mathbb{E}W = \mu^{-1}\mathbb{E}N$. Mais, comme ces lois ne semble déterminer même les ésperances, le point de vue moderne de lois de Little distributionelles semble incontournable.

7.6 Le problème de Dirichlet/première passage

Soit τ le temps de sortie d'un processus de Markov de son domaine $\mathcal{E}$, soit $c(x)$ defini sur le domaine et $b(x)$ defini sur le domaine complementaire ∂, et soit

$$f(x) = E_x[\int_0^\tau c(X(t))\ dt + \ b(X_\tau)]$$

Exercice 7.4 *Montrez, en conditionnant sur la position au moment $t+h$, pour h très petit, que $f(x)$ satisfait les équations de Dirichlet:*

$$\begin{aligned} (Gf)(x) + c(x) &= 0 \quad si\ \ x \in int(\mathcal{E}) \\ f(x) &= b(x)\ si\ \ x \in \partial \end{aligned}$$

Rémarque: Le problème de Dirichlet pour une chaîne discrète est obtenu en remplaçant le générateur G par son analogue discret $\frac{P-I}{1} = P - I$, ou P est la matrice de transition de la chaîne discrète (la démonstration directe de ce cas est pareille, est plus facile que celle du cas continu).

Théorème 7.2 *Etant donne un processus de Markov en temps continu X_u avec domaine $\mathcal{E}$ et generateur G, l'esperance du gain actualisé*

$$f(x) = E_x[e^{-\int_0^\tau r(X_u)du} b(X_\tau) + \int_0^\tau e^{-\int_0^t r(X_u)du} c(X(t))]\ dt$$

satisfait:

$$(Gf)(x) - r(x)f(x) + c(x) = 0$$

avec $f = b$ sur la frontière.

Note: Ces problèmes ont aussi des versions à durée finie T, où on s'interesse en:

$$f(x,T) = E_x[e^{-\int_0^{\min[\tau,T]} r(X_u)du} b(X_{\min[\tau,T]}) + \int_0^{\min[\tau,T]} e^{-\int_0^t r(X_u)du} c(X(t))]\ dt$$

Cette fois ci, le gain esperé f depend aussi du temps qui reste T, et on arrive a la solution d'une EDP a conditions initiales et de frontière:

$$\begin{aligned} \frac{\partial}{\partial T}\ f + \ (Gf)(x) - r(x)f(x) + c(x) &= 0 \quad \text{si}\ \ x \in int(\mathcal{E}) \\ f(x) &= b(x)\ \text{si}\ \ x \in \partial\ \text{ou}\ T = 0 \end{aligned}$$

7.7 La file M/M/1 et l'équation de Poisson

L'équation de Poisson. Pour X_t un processus de Markov ergodique avec générateur $\mathcal{G}$ et distribution stationnaire $\pi(dx)$, on a "normalement" pour chaque état initial x, par la "loi des grandes nombres" la convergence de la distribution empirique vers une distribution invariante $\pi(dx)$, i.e.

$$\int_0^t 1_{\{X_t \in [x,x+dx]\}} \frac{dt}{t} \to \pi(dx)$$

Cela entraîne, au moins pour des fonctions bornés $h(x)$, la convergence:

$$\gamma(h) = \lim_{t\to\infty} t^{-1}\mathbb{E}_x \int_0^t h(X_t)dt = \int_{-\infty}^{\infty} h(x)\pi(dx)$$

Parfois, on a aussi un résultat plus fort, l'existance de la limite "ajusté":

$$w(x) = \lim_{t\to\infty} \mathbb{E}_x \int_0^t (h(X_t) - \gamma(h))dt,$$

en quel cas $w(x)$ satisfait l'équation de Poisson

$$\mathcal{G}w(x) + (h(x) - \gamma) = 0 \quad \forall x \tag{28}$$

La fonction $w(x)$, appelé **"coût relatif"**, est determiné seulement "up to a constant". On peut fixer cette constante en choisissant un état arbitraire x_0 et en supposant que $w(x_0) = 0$. Ce choix particuler $\tilde{w}(x)$ vérifie

$$\tilde{w}(x) = \int_0^{T_0} (h(X_t) - \gamma(h))dt$$

et satisfait un système pareil au système:

$$\begin{aligned} \mathcal{G}c(x) + x &= 0 \quad \forall x \neq x_0 \\ c(x_0) &= 0 \end{aligned} \tag{29}$$

satisfait par le coût de vidage $c(x) = \mathbb{E}_x \int_0^{T_0} X_s ds$ jusqu'au temps de frappe T_0 de x_0, mais avec l'ajout de l'inconnue supplementaire γ. Il y a aussi une équation "de continuation" en 0, obtenue en appliquant (28) en 0, qui n'existait pas pour le systême (29) et qui détermine l'inconnue supplementaire.

Ainsi, l'equation de Poisson détermine au même temps le **"coût stationnaire"** γ et le "coût de vidage relatif" $\tilde{w}(x)$, qu'on denotera desormais par $w(x)$, quand le choix de x_0 est claire. C'est la son interêt, parce-que ces deux characteristiques sont importantes pour le contrôle optimal des files et des reseaux.

7.8 Processus quasi-naissance-et-mort/quasi-Toeplitz

L'étude des files d'attente Ph/Ph/s avec des temps d'arrivé ou départ type phase rammène a des problèmes avec générateur de type **quasi block-Toeplitz**, i.e. de structure:

$\boldsymbol{T}^{(0,0)}$	$\boldsymbol{T}^{(0,+)}$	0	0	0	0	...
$\boldsymbol{T}^{(1,-)}$	$\boldsymbol{T}^{(1,0)}$	$\boldsymbol{T}^{(1,+)}$	0	0	0	...
...	...	...	$\boldsymbol{T}^{(s-1,+)}$	...	...	...
$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{T}^{(s,-)}$	$\boldsymbol{T}^{(0)}$	$\boldsymbol{T}^{(+)}$	0	...
$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{T}^{(-)}$	$\boldsymbol{T}^{(0)}$	$\boldsymbol{T}^{(+)}$	0
...	...	...	...	...	...	...

ou $(\boldsymbol{T}^{(-)}, \boldsymbol{T}^{(0)}, \boldsymbol{T}^{(+)})$ sont des bloques constantes, en commençant d'un certain point. Par exemple, pour la file M/Ph/1 on trouve

-λ	$\lambda\,\boldsymbol{\beta}$	0	...	...	...
$\boldsymbol{b}$	$\boldsymbol{B}\ -\lambda\boldsymbol{I}$	$\lambda\boldsymbol{I}$	0	...	...
$\mathbf{0}$	$\boldsymbol{b}\boldsymbol{\beta}$	$\boldsymbol{B}\ -\lambda\boldsymbol{I}$	$\lambda\boldsymbol{I}$	$\mathbf{0}$	...
$\vdots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	...

et pour la file M/Ph/s on trouve:

-λ	$\lambda\,\boldsymbol{\beta}$	0	...	0	0
$\boldsymbol{b}$	$\boldsymbol{B}\ -\lambda\boldsymbol{I}_L$	$\lambda\left((\boldsymbol{\beta}\otimes\boldsymbol{I}_L)+(\boldsymbol{I}_L\otimes\boldsymbol{\beta})\right)$	...	0	0
$\mathbf{0}$	$(\boldsymbol{b}\otimes\boldsymbol{I}_L)+(\boldsymbol{I}_L\otimes\boldsymbol{b})$	$\boldsymbol{B}\oplus\boldsymbol{B}\ -\lambda\boldsymbol{I}_{L^2}$	$\lambda\boldsymbol{\beta}\oplus\boldsymbol{\beta}\oplus\boldsymbol{\beta}$	...	$\mathbf{0}$
...	...	$\boldsymbol{b}\oplus\boldsymbol{b}\oplus\boldsymbol{b}$	...	...	...
...	...	...	$\oplus_{i=1}^{s-1}\boldsymbol{B}_i\ -\lambda\boldsymbol{I}_{L^{s-1}}$	$\lambda\ \oplus_{i=1}^{s}\boldsymbol{\beta}_i$	$\mathbf{0}$
...	...	$\mathbf{0}$	$\oplus_{i=1}^{s}\boldsymbol{b}_i$	$\oplus_{i=1}^{s}\boldsymbol{B}_i\ -\lambda\boldsymbol{I}_{L^s}$	$\lambda\boldsymbol{I}_L$
...	...	...	...	$\oplus_{i=1}^{s}(\boldsymbol{b}_i\boldsymbol{\beta}_i)$	$\oplus_{i=1}^{s}\boldsymbol{B}_i$ -
...	...	...	...	...	$\oplus_{i=1}^{s}(\boldsymbol{b}$

Définition 7.3 *Un processus est appleé à naissances-et-morts en groupe quand son générateur est, au moins à une certaine distance des frontières, une matrice "quasi-Toeplitz", i.e. une matrice formé d'un nombre fini des "diagonales" constantes:*

$$G=\begin{pmatrix} -(\sum_i\Lambda_i) & \Lambda_1 & \Lambda_2 & 0 & 0 & \cdots & \\ A_1 & -(\sum_i\Lambda_i+\mu_1) & \Lambda_1 & \Lambda_2 & 0 & \cdots & \\ A_2 & A_1 & -(\sum_i\Lambda_i+A_1+A_2) & \Lambda_1 & \ddots & & \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & & & & \ddots & \ddots \end{pmatrix}$$

où Λ_i, A_i sont des matrices.

7.9 Exercices

1. Pour X_s un processus Poisson à taux λ, conditionné sur $X_0 = x$, et T le temps de la première arrivée, calculez la transformé de Laplace $\hat{f}_T[\delta] = \mathbb{E}e^{-\delta T}$, $\delta \geq 0$ et le coût total de stockage discompté

$$v_\delta(x) = \mathbb{E}_x \int_0^\infty e^{-\delta s} X_s ds$$

Indication: Utilisez la décomposition $X_s = x + Y_s$, ou Y_s est un processus Poisson conditionné sur $Y_0 = 0$. Soit $T^{(0)}$ le moment du premier saut. Alors, $Y_{T^{(0)}+s} = 1 + Y_s$.

2. Soit N_s le nombre de personnes en attente dans une file M/M/1/∞, soit $T = T_0$ le temps de vidage de la file (appellé aussi **pèriode occupèe/busy period** si $N_0 = 1$) et $T = \min[T_0, T_K]$ (où K peut-être la taille d'une salle d'attente).

   (a) Calculez **la distribution d'arrêt** $p(x) = P_x[N(T) = K]$, $q(x) = P_x[N(T) = 0]$. *Indication:* On peut conditionner après un petit interval dt, ou après le temps de la première transition.

   (b) **"Temps de vidage"**. Montrez que si $\lambda < \mu$, $\delta > 0$, et $K = \infty$, alors

      i. Le **temps de vidage esperé** $t(x) = \mathbb{E}_x T_0$ est de la forme $t(x) = \tau\ x$; déterminez τ.

      ii. Montrez que la **transformée de Laplace du temps de vidage** $d(x) = \mathbb{E}_x e^{-\delta T_0}$ est de la forme $d(x) = z^x$; déterminez z. Donnez des intérpretations probabilistes pour z, τ.

      iii. Calculez aussi la transformée de Laplace du "temps de repos" -"idle-time".

3. **Quelques coûts interéssants pour la file M/M/1.**

   (a) **"Coût stationnaire" de la file M/M/1**. Calculez le côut moyen stationnaire $\gamma = \mathbb{E}_{ss} X_t$ de la file M/M/1.

   (b) **"Coût de vidage esperé" la file M/M/1**. Obtenez et resolvez l'équation de recurrence pour le coût de vidage $c(x) = \mathbb{E}_x \int_0^T X_s ds$, ou $T = T_0$. Montrez que

$$c(x) = t(x)(\frac{x+1}{2} + \gamma)$$

   ou $\gamma = \mathbb{E}_{ss} X_t$ est le côut moyen stationnaire de la file M/M/1. Donnez une interpretation probabiliste de la partie $w(x) = t(x)\frac{x+1}{2} = c(x) - \gamma t(x)$.

   (c) Obtenez et resolvez l'équation de recurrence pour le **coût de vidage discompté**

$$c_\delta(x) = \mathbb{E}_x \int_0^T e^{-\delta s} X_s ds$$

   Trouvez aussi la limite $b(x)$ ("biais") de la dernière fonction quand $\delta \to 0$. Est-ce que la limite commute avec l'ésperance?

   (d) Calculez pour la file M/M/1 le "coût relatif" au coût stationnaire $c_r(x)$, i.e. la solution de l'equation de Poisson

$$\mathcal{G} c_r(x) + (x - \gamma) = 0 \quad \forall x, \quad c_r(0) = 0, \quad c_r(x)(\text{"temperèe à } \infty\text{"})$$

   et comparez avec la limite precedente $b(x)$.

(e) Calculez la **mesure d'occupation**

$$m_x(k) = \mathbb{E}_x \int_0^{T_0} 1_{\{X_t=k\}} dt$$

Indication: Obtenez d'abord la fonction generatrice $M(x) = \sum_k z^k m_x(k) = \mathbb{E}_x \int_0^{T_0} z^{X_t} dt$.

(f) Calculez la **mesure d'occupation rélative** $m_{r,x}(k)$, en commençant par la solution de l'equation de Poisson

$$\mathcal{G}M_r(x) + (z^x - \mathbb{E}_{SS} z^{X_t}) = 0 \quad \forall x, \quad M_r(0) = 0, \quad M_r(x)(\text{"temperèe à } \infty\text{"})$$

et comparez avec le résultat precedent.

(g) Calculez, en conditionnant après un petit interval dt, ou après le temps de la première arrivée en 0, le **"coût total discompté** $v_\delta(x) = \mathbb{E}_x \int_0^\infty e^{-\delta s} X_s ds$, et la limite $\lim_{\delta \to 0} v_\delta(x)$.

4. a) Montrez que les processus "discretisés" obtenues en regardant la file d'attente M/G/1 aux temps des départs et la file G/M/1 aux temps d'arrivé sont des chaînes de Markov en temps discret, en spécifiant leur matrices de transition.

   b) Quelles sont les matrices de transition des chaîne s discretisés des files $M/G/1/K$ et $G/M/1/K$?

5. Repetez les questions du problème 3) pour une file M/M/∞.

6. Repetez les questions du problème 3) pour une file Brownien $B_{\mu,\sigma^2}(t)$(directement, ou en choisissant des parametres $\lambda, \mu \to \frac{1}{2}$ tel que la file M/M/1 approxime un mouvement Brownien).

7. Repetez les questions du problème 3) pour un réseaux des deux files M/M/1 en tandème.

8. Calculez la distribution stationnaire de la file M/Ph/∞.

9. Calculez la distribution stationnaire de la file M/Ph/1, et la probabilité $\pi(s)$ que les nouveaux arrivés ne doivent pas attendre.

10. (***) Calculez pour la file M/Ph/s la probabilité $\pi(s)$ que les nouveaux arrivés ne doivent pas attendre.

11. Trouvez les fonction génératrices et transformées de Laplace de N, Q, T, W pour la file M/G/1 (la formule de Pollaczek-Khyncin).

12. Calculez la transformée de Laplace de la pèriode occupée d'une file M/Ph/1. Vérifiez cette formule avec le cas M/M/1.

13. Montrez que la transformée de Laplace de la pèriode occupée d'une file M/G/1 satisfait **l'équation de Kendall**:

$$z[s] = \hat{b}[s + \lambda - \lambda z[s]] \tag{30}$$

14. Est ce que la file "B(atch)"M/Ph/∞ a une distribution de forme produit?

Solutions:

1. Le temps de la première arrivée étant exponentiel, $\mathbb{E}e^{-\delta T} = \frac{\lambda}{\lambda+\delta}$. Conditionnant sur le temps de la première arrivée, on trouve d'abord $v(0) = \frac{\lambda}{\lambda+\delta}(\frac{1}{\delta} + v(0))$ et alors: $v(0) = \frac{\lambda}{\delta^2}$. En suite, $v(x) = \frac{\delta}{\lambda+\delta}\frac{1}{\delta} + \frac{\lambda}{\lambda+\delta}v(x+1)$, ...

2. (a) Soit $(\mathcal{G}p)_x = \lambda(\ p_{x+1} - p_x) + \mu\ (p_{x-1} - p_x)$ En conditionnant apres un petit interval dt, on trouve:

   i. $p_x = P_x[X(T) = K]$ satisfait:

$$\begin{aligned} (\mathcal{G}p)_x &= 0 \quad \text{for any } 1 \le x \le K-1 \\ p_K &= 1 \\ p_0 &= 0 \end{aligned}$$

   Comme c'est une equation homogêne, et les racines de l'equation auxiliare sont $\frac{\mu}{\lambda}$ and 1, on obtient $p_x = A_1 + A_2(\frac{\mu}{\lambda})^x = \frac{1-(\frac{\mu}{\lambda})^x}{1-(\frac{\mu}{\lambda})^K}$.

   ii. $d(x) = \mathbb{E}_x e^{-\delta T}$ satisfait:

$$\begin{aligned} (\mathcal{G}d)(x) - \delta d(x) &= 0 \\ d(0) &= 0 \\ d(\infty) &\le 1 \end{aligned}$$

   Comme c'est une equation homogêne, et l'equation auxiliare

$$\lambda a^2 - (\lambda + \mu + \delta)a + \mu = 0 \Longleftrightarrow \mu(a^{-1} - 1) + \lambda(a - 1) = (1 - a)\frac{\mu - a\lambda}{a} = \delta$$

   a exactement une racine z qui n'est pas plus grande que 1, on obtient $p_x = z^x$. Par exemple, avec $\delta = 0$ on obtient $z = 1$.

   iii. $t_x = E_x[T]$ (temps d'arrêt esperé) satisfait une equation nonhomogène

$$\begin{aligned} (\mathcal{G}t)_x + 1 &= 0 \\ t_0 &= 0 \\ t_x &\quad \text{accroit non-exponentiellement quand } x \to \infty \end{aligned}$$

   Comme c'est une equation nonhomogêne, on obtient $t_x = p_x + \tilde{t}_x$, ou $\tilde{t}_x = cx$. La solution particulière est $\tilde{t}_x = \frac{x}{\mu-\lambda}$, la condition a ∞ annule le terme

$A_2(\mu/\lambda)^x$ et la condition initiale annule A_1, et alors la solution de l'equation nonhomogene est $t(x) = \frac{x}{\mu-\lambda}$.
τ, z sont l'ésperance et la transformé de Laplace d'une periode de decroisement par 1 de la file ("busy period"). On peut aussi interpreter z comme la valeur actualise (par rapport au taux δ d'une unite monetaire qu'on recevra apres que la file decroise par 1.
Remarque: La distribution de la periode de decroisement par 1 de la file ("busy period") est tres importante justement a cause de la structure additive/multiplicative signale ci dessu.

3. (a) En utilisant $\gamma = \sum_i i(1-\rho)\rho^i$ où $\rho = \lambda/\mu$, on trouve:

$$\gamma = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

(b) Pour $c(x) = E_x[\int_0^T X(t)dt]$ (coût totale de stockage) on trouve

$$\begin{aligned} (\mathcal{G}c)(x) + x &= 0 \\ c(0) &= 0 \\ c(x) & \quad \text{accroit non-exponentiellement quand } x \to \infty \end{aligned}$$

Comme $\mathcal{G}x = \lambda - \mu$ et $\mathcal{G}x^2 = 2x(\lambda-\mu) + \lambda + \mu$, on trouve par la methode des coefficients indeterminés que la solution est $c(x) = \frac{x^2}{2(\mu-\lambda)} + \frac{x(\mu+\lambda)}{2(\mu-\lambda)^2}$
On trouve que $c(x) = c_r(x) + \gamma t(x)$. **Interpretation:** La première partie, $c_r(x) = t(x)(\frac{x+1}{2})$ est la solution "fluide", qui utilise precisement la moyenne arithmetique des valeurs $(x, x-1, ..., 1)$, comme si avant le vidage on aurait visité seulement ces valeurs, pour des durées de temps egaux. La deuxième partie sugère un "remplacement de la valeur moyenne a long terme 0 ("vue de loin") dans la solution fluide par la valeur moyenne réelle a long terme γ.

(c) Pour $c_\delta(x) = \mathbb{E}_x \int_0^T e^{-\delta s} X_s ds$ on trouve

$$\begin{aligned} (\mathcal{G}c)(x) - \delta c + x &= 0 \\ c(0) &= 0 \\ c(x) & \quad \text{accroit non-exponentiellement quand } x \to \infty \end{aligned}$$

avec la solution $c_\delta(x) = \frac{x}{\delta} + \frac{\lambda-\mu}{\delta^2}(1-z^x)$, ou $z = \mathbb{E}_1 e^{-\delta T_0}$.
Pour la limite, on utilise $\lim_{\delta\to 0} \frac{1-z}{\delta} = \frac{1}{\mu-\lambda}$, $\lim_{\delta\to 0} \frac{1-z^i}{\delta} = \frac{i}{\mu-\lambda}$. Alors

$$\begin{aligned} b(x) = \lim c_\delta(x) &= \frac{x + (\lambda-\mu)\frac{1-z}{\delta}\frac{1-z^x}{1-z}}{\delta} \\ &\approx \frac{\sum_{i=0}^{x-1}(1-z^i))}{\delta} = \frac{x(x-1)}{2(\mu-\lambda)} \end{aligned}$$

(d) Pour $v_\delta(x) = \mathbb{E}_x \int_0^\infty e^{-\delta s} X_s ds$ on trouve

$$\begin{aligned}(\mathcal{G}v)(x) - \delta v(x) + x &= 0\\ v(0) &= \frac{\lambda}{\lambda+\delta} v(1)\\ v(x) &\quad \text{accroit non-exponentiellement quand } x \to \infty\end{aligned}$$

Tenant compte de $v_\delta(x) = c_\delta(x) + z^x v(0)$, on trouve que $v(0) = \frac{\lambda(1-z)}{\delta^2}$ et finalement

$$\begin{aligned}v(x) &= \frac{x}{\delta} + \frac{\lambda-\mu}{\delta^2} + z^x \frac{\mu-\lambda z}{\delta^2} = \frac{x}{\delta} + \frac{\lambda-\mu}{\delta^2} + \frac{z^{(x+1)}}{\delta(1-z)}\\ &= \frac{x}{\delta} + \frac{\lambda-\mu}{\delta^2} + z^x \frac{\mu-\lambda z}{\delta^2} = \delta^{-1}(\sum_{i=1}^{x}(1-z^i) - 1) + \frac{\lambda-\mu}{\delta^2} + \frac{1}{\delta(1-z)}\\ &= \delta^{-1}(\sum_{i=1}^{x}(1-z^i)) + \frac{\lambda-\mu}{\delta^2} + \frac{\delta z}{\delta^2(1-z)}\\ &= w(x)???\end{aligned}$$

(e) Comme en 1a), en utilisant aussi l'équation "de continuation" en 0 $\lambda c_r(1) = \gamma$, on trouve $c_r(x) = \frac{x(x+1)}{2(\mu-\lambda)}$. Ainsi, $\boxed{c_r(x) = b(x+1)}$ et on trouve aussi $\boxed{c_r(x) = c(x) - \gamma t(x)}$ (on peut calculer γ plus directement, en partant de proprietes de $c_r(x)$, comme le minimum a $x = -1/2???$)

(f) Pour $M(x) = \mathbb{E}_x \int_0^T z^{X_s} ds, z \le 1$ on trouve:

$$\begin{aligned}(\mathcal{G}M)(x) + z^x &= 0\\ M(0) &= 0\\ M(x) &\quad \text{accroit non-exponentiellement quand } x \to \infty\end{aligned}$$

Comme $\mathcal{G}z^x = z^x(\lambda(z-1) + \mu(z^{-1}-1)) := z^x\,\kappa(z)$ où $\kappa(z) := (\lambda(z-1) + \mu(z^{-1}-1)) = (\mu z^{-1} - \lambda)(1-z) = z^{-1}(\mu - \lambda z)(1-z)$, on trouve que:

$$\begin{aligned}M(x) = \frac{1-z^x}{\kappa(z)} = \frac{z-z^{x+1}}{(\mu-\lambda z)(1-z)} &= \frac{z-z^{x+1}}{\mu-\lambda}(\frac{1}{1-z} - \frac{\lambda}{\mu-\lambda z})\\ = \frac{z-z^{x+1}}{\mu-\lambda}(\frac{1}{1-z} - \frac{\rho}{1-\rho z}) &= \frac{z-z^{x+1}}{\mu-\lambda}\sum_{k=0}^{\infty} z^k(1-\rho^{k+1})\\ &= \sum_{k=1}^{x} z^k \frac{1-\rho^k}{\mu-\lambda} + \sum_{k=x+1}^{\infty} z^k \frac{\rho^{k-x} - \rho^k}{\mu-\lambda}\end{aligned}$$

En conclusion, en inversant la transformé on voit que la mesure d'occupation est croissante et approximativement uniforme $m_x(k) \approx \frac{1-\rho^k}{\mu-\lambda}$ sur $\{1, x\}$, et la difference entre une géometrique deplacé en $x+1$ et une géometrique (donc approximativement 0) sur $\{x+1, \infty\}$. On a obtenu donc ici et un rafinement exact de l'approximation fluide!

4. **Sol b):** Pour la file $M/G/1/K$, posons $p_k = \mathbb{P}\{k \text{ arrivés pendant un service }\}$ et $P_k = \sum_{i=k}^{\infty} p_i$ pour le premier cas.

$$G = \begin{pmatrix} p_0 & p_1 & p_2 & \cdots & \cdots & P_{K+1} \\ p_0 & p_1 & p_2 & \cdots & \cdots & P_{K+1} \\ 0 & p_0 & p_1 & p_2 & \cdots & P_K \\ \vdots & \ddots & \ddots & \ddots & \ddots & P_{K-1} \\ \vdots & \ddots & \ddots & \ddots & p_1 & \vdots \\ 0 & 0 & \cdots & 0 & p_0 & P_1 \end{pmatrix}$$

Pour la file $G/M/1/K$, posons $p_k = \mathbb{P}\{k \text{ services entre deux arrivés }\}$ pour le deuxième cas et $P_k = \sum_{i=k}^{\infty} p_i$ pour le premier cas...

8 Exit/first passage problems

8.1 Actuarial and queueing origins

The Cramer-Lundberg, renewal, and jump-diffusion risk models. First passage theory, an extension of the classical gambler's ruin problem, is nowadays concerned with exit problems for "reserves processes" of the form

$$X_t := x + \mu t - \sum_{i=1}^{N(t)} Z_i + Y_t. \tag{31}$$

In the actuarial science interpretation, x denotes the initial reserve of an insurance company, μ is the premium rate per unit of time, and the "claims/catastrophes" Z_i are i.i.d. random variables with distribution and density concentrated on $(0, \infty)$, and denoted by $F(x), f(x)$, respectively. The "cumulated claims" process $S(t) = \sum_{i=1}^{N(t)} Z_i$ involves an independent renewal process of claim arrivals

$$N(t) = \inf\{k \in \mathbb{N} : \sum_{i=1}^{k} a_i \leq t\},$$

where a_i denote i.i.d. nonnegative inter-arrival times. We will denote by λ and μ the reciprocals of the means of the claims and inter-arrival times, respectively. Finally, the process Y_t denotes an independent process modeling fluctuations of smaller order than the catastrophes, and is typically taken to be a Brownian motion.

In the case when a_i are exponential ($N(t)$ being thus a Poisson process of intensity λ) and in the absence of the "perturbation" Y_t, (31)) is known as the classical Cramer-Lundberg model, in honor of the pioneers who laid the foundations of actuarial science, by their investigations at the beginning of the century –see [**?**], [**?**]. In the presence of the small fluctuations part Y_t we have the "perturbed model", while the case of general interclaim times a_i is known as the renewal, or Sparre Andersen model.

First-passage problems. Let T_y^+, T_y denote the times of exit/ first passage out of a half-line:

$$T_y^+ := \inf\{t \geq 0 : X_t \geq y\}$$

and

$$T_y := \inf\{t \geq 0 : X_t < y\}$$

More generally, we will denote by $T_A := \inf\{t \geq 0 : X_t \notin A\}$ the exit time from a set A. When A is an interval, T_A is known as a two-sided exit time (distinguishing it from the one-sided exit out of a half-line).

The first passage time over 0,

$$\tau := T_{(-1,0)} = \inf\{t \geq 0 : X_t < 0\}$$

is also called **ruin time**.

Classical risk theory attempts to compute the finite-time first-passage probabilities

$$\psi(t, x) = P_x[\tau \leq t] \tag{32}$$

or survival probabilities $\overline{\psi}(t, x) = 1 - \psi(t, x)$.

Under the markovian model (46), taking into account the respective boundary conditions, we find that these functions satisfy respectively:

$$\tilde{\mathcal{G}}\ \psi(x) - (\lambda + q)\psi(x) + \lambda \bar{B}(x) = \frac{\partial}{\partial t}\psi(t, x),\ \psi(t, x) = 1_{x<0},\ \psi(t, \infty) = 0$$
$$\tilde{\mathcal{G}}\overline{\psi}(x) - (\lambda + q)\overline{\psi}(x) = \frac{\partial}{\partial t}\overline{\psi}(t, x), \quad \overline{\psi}(t, x) = 0, x < 0, \quad \overline{\psi}(t, \infty) = 1$$

where $\tilde{\mathcal{G}}h(x) = [(c(x)\frac{\partial}{\partial x} + d(x)\frac{\partial^2}{\partial x^2}]\ h(x) + \lambda \int_0^x h(x - z)F(dz)$.

These partial differential equations are seldom tractable and the only analytic results for finite-time ruin probabilities are when X_t is Brownian motion (and the first-passage time is inverse Gaussian), and for the model with exponential claims -see [**?**].

More analytical results are available for "ultimate" or "perpetual" first passage/ruin probabilities

$$\psi(x) = P_x[T < \text{ı}] \tag{33}$$

or for the "killed" probabilities of passage before an independent exponential horizon $\mathcal{E}(q)$ of rate q:

$$\psi_q(x) = P_x[T \leq \mathcal{E}(q)] = \int_0^\infty q\mathrm{e}^{-qt} P_x[T \leq t]dt = \int_0^\infty \mathrm{e}^{-qt}\psi(x, \mathrm{d}t) = E_x\mathrm{e}^{-qT} \tag{34}$$

For the classical Cramer-Lundberg process for example, the killed ruin and survival probabilities satisfy respectively:

$$\tilde{\mathcal{G}}\psi_q(x) - q\psi_q(x) + \lambda\bar{P}(x) = 0$$
$$\tilde{\mathcal{G}}\overline{\psi}_q(x) - q\overline{\psi}_q(x) + q = 0$$

which after taking into account the boundary conditions and replacing λ by its scaled version $\tilde{\lambda} = \frac{\lambda}{\mu}$, may be put into the form:

$$\overline{\psi}(x) - \overline{\psi}(0) = \tilde{\lambda} \int_0^x \overline{\psi}(x-y)\bar{P}(y)dy, x \geq 0 \tag{35}$$

Notes: 1) The killed downwards passage probabilities are the analog of the $q-$ harmonic positive decreasing function for diffusions, i.e. solutions of Sturm-Liouville equations $G\psi_q(x) - q\psi_q(x) = 0$; however, the multiplicative structure implied by the smooth passage is lost.

2) Since this quantity is also the Laplace-Carson transform in time of $\psi(t,x)$, its knowledge allows recovering the finite time ruin probabilities numerically, by inversion of the Laplace transform.

3) Throughout, perpetual/killed versions will be indicated by suppressing the variable t/adding a subscript q.

4) The solution of the Volterra equation (35) which may be written also as:

$$\overline{\psi}(x) * \left(\delta(x) - \tilde{\lambda}(x_+^0 - P(x))\right) = \overline{\psi}(0)\ x_+^0$$

may be developed in a Neumann series (involving the convolution powers of the kernel $k(x) = \tilde{\lambda}(x_+^0 - P(x))$, yielding:

$$\overline{\psi}(x) = \sum_{n=0}^{\infty} \overline{\psi}(0)\ \psi(0)^n F_e(x)^{*n} \tag{36}$$

$$\psi(x) = \sum_{n=1}^{\infty} \overline{\psi}(0)\ \psi(0)^n \bar{F}_e(x)^{*n} \tag{37}$$

where $\rho := \psi(0) = \frac{\lambda}{\mu\mu}$, and where $\bar{F}_e(x)$ is the excess claim distribution[9]. This expansion was first obtained by Dubordieu(1952), Benes(1957) and Kendall(1957), and it is known as the ladder decomposition due to its probabilistic interpretation, and also as the Pollaczek-Khinchin formula, even though these researchers only provided the Laplace transform of the series.

Two alternative forms of the DBK series –see Shiu –[**?**] are Takacs's formulas:

$$\overline{\psi}(x)/\overline{\psi}(0) = \sum_{n=0}^{\infty} \frac{(-\tilde{\lambda})^n}{n!} E\left[(x - S_n)_+^n\ \mathrm{e}^{\tilde{\lambda}(x-S_n)}\right] \tag{38}$$

$$= \mathrm{e}^{\tilde{\lambda}x} + \int_0^x \tilde{f}_{y-x}(y)dy \tag{39}$$

$$\psi(x)/\overline{\psi}(0) = \sum_{n=1}^{\infty} \frac{(\tilde{\lambda})^n}{n!} E\left[(S_n - x)_+^n\ \mathrm{e}^{\tilde{\lambda}(x-S_n)}\right] \tag{40}$$

$$= \int_x^{\infty} f_{y-x}(y)dy \tag{41}$$

[9]when $\mu = 1$, this coincides with the density of the stationary waiting time of the M/G/1 queue

where $f_t(y) = \sum_{n=1}^{\infty} \frac{(\tilde{\lambda}t)^n}{n!} p^{*n}(y)$ is the density of the compound Poisson rv $\sum_{i=1}^{N_t} Z_i$ and $\tilde{f}_t(y)$ is its analytic continuation to negative values of t. Note that the sum of the two formulas reveals an interesting nontrivial identity for $1/\overline{\psi}(0)$.

These formulas show that infinite horizon ruin probabilities are deducible from the semigroup transition kernel. The latter is however rarely available analytically.

Exemple 8.1 *With exponential claim sizes of intensity μ, the ultimate ruin probability is*[10]*:*

$$\psi(x) = C\mathrm{e}^{-\gamma x} \text{ where } \gamma = \mu - \lambda/\mu > 0,\ C = \frac{\lambda}{\mu\mu} < 1.$$

One way to obtain this result (and its phase-type generalization) is from the Dubordieu-Benes-Kendall series. Another, much easier way in this case, is to look directly for exponential solutions of the corresponding harmonic equation.

Notes: 1) When the jumps have positive integer sizes, Takacs's formula (38) reduces to a finite sum:

$$\overline{\psi}(x)/\overline{\psi}(0) = \mathrm{e}^{\tilde{\lambda}x} + \sum_{k=1}^{\lfloor x \rfloor} \mathrm{e}^{\tilde{\lambda}(x-k)} \sum_{n=1}^{k} \frac{P[S_n = k]\ [\tilde{\lambda}(k-x)]^n}{n!}, \tag{42}$$

as noted by Shiu [**?**].

2) The further particular case of deterministic claims of size 1 can be derived directly from the ladder expansion series. The excess claim distribution is uniform on $[0, 1]$. The c.d.f. of the n-th convolution is given by [11]:

$$n!F_n(x) = \sum_{k=0}^{\lfloor x \rfloor} (-1)^k \binom{n}{k} (x-k)^n$$

leading finally to:

$$\frac{\overline{\psi}(x)}{1-\tilde{\lambda}} = \sum_{n=0}^{\infty} \tilde{\lambda}^{n-1} F_n(x) = \sum_{k=0}^{\lfloor x \rfloor} \frac{(-\tilde{\lambda}(x-k))^k}{k!} e^{\tilde{\lambda}(x-k)}$$

which is of course a particular case of Shiu's formula (42).

In the more general case of the M/D/1 queue, the corresponding waiting time distribution formula becomes

$$P[W \le x] = (1-\rho) \sum_{k=0}^{\lfloor x \rfloor} \frac{(-\lambda(x-kD))^k}{k!} e^{\lambda(x-kD)}$$

[10] Explicit formulas involving matrix exponentials are always available when the claims have rational Laplace transform.

[11] this seems to suggest that a formula relating the convolutions of an initial density $f(z)$ and its complementary cumulative distribution $\int_x^{\infty} f(z)dz$ is possible

This formula (due to Erlang (1909) and generalized to M/D/c by Crommelin) is not –as noted in Feller– of real practical interest, due to the cancelation of huge exponentials with positive exponents. See Franx [**?**] for an alternative expression and simple derivation.

The joint ruin-severity of ruin and the integrated survival distributions. More generally, the first passage problem below 0 consists in computing the related "ruin-severity of ruin" and "integrated survival" joint moment generating functions:

$$\Psi(x,\theta) = \Psi_q(x,\theta) = E_x[\mathrm{e}^{-qT+\theta X_T}], \quad i(x,\theta) = i_q(x,\theta) = E_x \int_0^T \mathrm{e}^{-qs+\theta X_s} ds \tag{43}$$

for such values $q > 0$ and θ such that they are finite.

Phase-type distributions. Especially convenient in applications are **phase-type models**, for which either the distribution of Z_i or of a_i belong to the class of phase-type distributions generalizing the exponential.

Définition 8.1 *A distribution F on $(0,\infty)$ is phase-type if it is the distribution of the absorbtion time ζ of a finite continuous time Markov process $J = \{J_t\}_{t\geq 0}$ with one state Δ absorbing and the remaining ones $1,\ldots,m$ transient. That is, $F(t) = \mathbb{P}(\zeta \leq t)$ where $\zeta = \inf\{s > 0 : J_s = \Delta\}$.*

Let m, $\boldsymbol{A}$, denote the number of transient states and the restriction of the full intensity matrix to the transient states, and let $\boldsymbol{\alpha} = (\alpha_1 \ldots \alpha_m)$ where $\alpha_i = \mathbb{P}(J_0 = i)$ denote the initial probability (row) vector. For any $i = 1,\ldots,m$, let a_i be the intensity of a transition $i \to \Delta$ and write $\boldsymbol{a} = (a_1 \ldots a_m)'$ for the (column) vector of such intensities. Further, let $\mathbf{1}_m$ denote a $m \times 1$ column vector of ones. It follows that the complementary cumulative distribution function $\bar{F}(x) = P[\zeta > x]$ is given by $\bar{F}(x) = \boldsymbol{\alpha}\mathrm{e}^{\boldsymbol{A}x}\mathbf{1}_m$, and the density and Laplace transforms are respectively $f(x) = \boldsymbol{\alpha}\mathrm{e}^{\boldsymbol{A}x}\boldsymbol{a}$, $f^*(s) = \boldsymbol{\alpha}(\boldsymbol{I} - s\boldsymbol{A})^{-1}\boldsymbol{a}$, , with $\boldsymbol{a} = -\boldsymbol{A}\,\mathbf{1}_m$.

The excess complementary cumulative distribution function is given by $\bar{F}_e(x) = \boldsymbol{\eta}\mathrm{e}^{\boldsymbol{A}x}\mathbf{1}_m$, where $\boldsymbol{\eta} = \boldsymbol{\alpha}(-\boldsymbol{A})^{-1}/m_1$ and $m_1 =< \boldsymbol{\alpha}(-\boldsymbol{A})^{-1}, \mathbf{1}_m >$ is the first moment.

Phase-type distributions include and generalize exponential distributions in series and/or parallel and form a dense class in the set of all distributions on $(0,\infty)$. Much of the applicability of the class comes from the probabilistic property that the overshoot distributions $F(x+y)/(1-F(x))$ belong to a finite vector space (in fact, the overshoot distribution is again phase–type with the same m and $\boldsymbol{T}$ but α_i replaced by $\mathbb{P}(J_x = i|\zeta > x)$). This provides matrix analogues of many formulas for the exponential distribution ($m = 1$) based upon the memoryless property. Other advantages, stemming from the rationality of the Laplace transforms, are the facts that the integral operators inherent to the study of jump models may be avoided, and that the Wiener-Hopf factorization necessary for inverting Laplace transforms may be replaced by matrix manipulations. See for example [**?**, **?**] for surveys.

8.2 Levy processes

The reserves process (31) with Poisson arrivals and Y_t Brownian motion is an example of a **Levy process without upward jumps** (spectrally negative), and in fact most of the results about it are also valid in this greater generality.

While this family of processes includes a bewildering variety of instances, with quite different path properties, it does enjoy a unifying element. Typically, results about Levy process admit a unified expression valid for the whole family, in terms of the moment generating function, also called **cumulant exponent/symbol/kernel** $\kappa(\theta)$:

$$\kappa(\theta) := t^{-1} \log\big(E_0\big[\mathrm{e}^{\theta X(t)}\big]\big),$$

well defined on some maximal domain

$$\Theta = \{\theta \in \mathbb{R} : \kappa(\theta) < 1\}$$

whose interior will be denoted by Θ^o. Let $[l, r]$ denote the closure of Θ. The map $s \mapsto \kappa(s)$ restricted to this interval of the real line is a convex function. We will always assume that the $0 \in \Theta^o$, in which case κ may be extended analytically to the strip $\Re(s) \in (l, r)$ in the complex plane.

Exemple 8.2 *The cumulant exponent of the Cramer-Lundberg model* $X(t) = X(0) + \sigma W_t + \mu\, t - \sum_{i=1}^{N_t} Z_i$ *perturbed by an independent standard Brownian motion* W *is:*

$$\kappa(\theta) = \frac{\sigma^2\theta^2}{2} + \mu\theta + \lambda(\hat{f}[\theta] - 1)$$

This is a particular case of the Levy-Khincin decomposition. Note that the exponent is a sum of three parts: one involving the distribution of the jumps, and the other two indicating respectively the presence of Brownian motion and of a deterministic drift.

Levy processes are Markovian. The generator of a Levy process is related formally to its symbol via $G = \kappa(D)$, where D is the first derivative (and where this expression may be interpreted in the sense of pseudo-differential operators).

By the Levy-Khinchin formula, the generator G is a sum of three operators; two indicate respectively the presence of Brownian motion and of a deterministic drift, and the third, the convolution operator with the density f, corresponds to the possibility of immediate jumps, and may also be written formally as $\hat{f}(D)$. When the Laplace transform $\hat{f}(D) = \frac{p(D)}{q(D)}$ is rational, this may be used to convert the integro-differential equations to ODE's, simply by multiplying with the denominator $q(D)$!

One way the symbol intervenes is in the computation of "$q-$ harmonic/Sturm-Liouville" functions which satisfy the "Sturm-Liouville equation"

$$\boxed{Gf(x) - qf(x) = 0}, \quad q > 0 \tag{44}$$

where G is the differential operator associated to the Markovian process $X(t)$. These functions are known to be of exponential form $\mathrm{e}^{\theta x}$ and since

$$G\mathrm{e}^{\theta x} = \kappa(\theta)\, \mathrm{e}^{\theta x},$$

finding them reduces to solving the algebraic equation

$$\kappa(\theta) - q = 0, \quad \forall q \geq 0 \tag{45}$$

which is also known as "Lundberg's equation". Its solutions, and especially the inverses $q_+ = \Phi(q)$, $q_- = \Phi_-(q)$ of $\kappa(\theta)$, defined respectively as the largest/smallest solutions within the strip $[l, r]$ of analiticity of the symbol, play an important role in this theory. Note that formally, The Cramer-Lundberg equation is a "scalar version" of the Sturm-Liouville equation, since formally $G = \kappa(D)$.

First passage problems for general Levy processes may be solved in principle via the "Wiener-Hopf-Rogozin factorization"

$$\frac{q}{q - \kappa(\theta)} = E e^{\theta S_{\mathcal{E}(q)}} \; E e^{\theta I_{\mathcal{E}(q)}}$$

where S, I denote respectively the supremum and infimum of our process and the RHS may be computed by factoring the LHS into two functions which are analytic over $\Re(\theta) \geq 0$ and $\Re(\theta) \leq 0$, respectively. In the spectrally one-sided case, the factorization reduces simply to the determination of the **unique** nonnegative/nonpositive root, and this uniqueness simplifies considerably the analysis [12].

8.3 Wong-Pearson diffusions

We will be interested below in more general state dependent Markovian risk processes given by a jump-diffusion with negative jumps:

$$\begin{aligned} dX_t &= c(X_t)dt - dS(t) + \sigma(X_t)dB(t) \\ S(t) &= \sum_{i=1}^{N(t)} Z_i \end{aligned} \tag{46}$$

where $c(x)$ is a state-dependent premium rate, $B(t)$ is standard Brownian motion (introducing a source of volatility $\sigma(X_t)dB(t)$ in the premiums accrual), and $S(t)$ is a compound Poisson process.

The "jump diffusion " process X_t is Markovian, with infinitesimal generator :

$$G\, h(x) := [\frac{\sigma(x)^2}{2}\frac{\partial^2}{\partial x^2} + c(x)\,\frac{\partial}{\partial x}]\; h(x) + \int_0^\infty [h(x-z) - h(x)]\, \nu(dz) \tag{47}$$

where $\nu(dz) = \lambda\, f(z)dz$ is the Levy measure.

In applications, the jump-diffusion model (46) may be viewed as a hypothesis on the existence of parameters so that the "residual process" $dB(t)$ is white noise. For this to be practical, it is useful of course to restrict oneself to models with few parameters.

One such class is obtained by assuming no jumps and by supposing the "drift" and "volatility" parameters $c(x)$ and $d(x) := \frac{\sigma^2(x)}{2}$ to be first and second order polynomials,

[12] More generally, with phase-type jumps in one direction, the WH factorization reduces to finding a finite number of nonnegative/nonpositive roots; furthermore, this may be achieved via matrix computations, by representing the symbol in matrix-exponential form.

respectively. As noted by Wong [**?**], for this class the first order equation for the speed measure

$$\frac{\mathbf{W}'(x)}{\mathbf{W}(x)} = \frac{c(x) - d'(x)}{d(x)} = \frac{b_0 + b_1 x}{d_0 + d_1 x + d_2 x^2} \tag{48}$$

(which yields the stationary distribution in the ergodic case) is precisely the famous "Pearson equation" introduced by K. Pearson (1914) in [**?**][13].

9 First passage of affine spectrally negative jump-diffusions

9.1 Integro-differential equations for penalty-reward functions of stopped "state dependent Levy processes"

To treat simultaneously the ruin, survival, and other related penalty-reward functions, assume given two monotone, bounded functions π, P, which represent respectively:

- The penalty at ruin $\pi(X_T)$ with deficit X_T, $\quad \pi : \mathbb{R}^- \to \mathbb{R}^-$
- The reward on survival after t years: $P(X_t)$, $P : \mathbb{R}^+ \to \mathbb{R}^+$.

It may be further useful to distinguish between the events $X_T = 0$ and $X_T < 0$. Let

$$v(t,x) = E_x \left(P(X_t)\, 1_{\{T \geq t\}} + \pi\left(X_T\right)\, 1_{\{T < t, X_T < 0\}} + \pi(0)\, 1_{\{T < t, X_T = 0\}} \right) \tag{49}$$

denote the expected penalty/reward function.

In the case $t = \infty$, this becomes the "perpetual/ultimate" penalty/reward:

$$v(x) = P(\infty)\, \mathbb{P}_x\{T = \infty\} + \mathbb{E}_x \pi\left(X_T\right)\, 1_{\{T < \infty\}}$$

Examples: Some particular cases of interest are the survival probability within t years

$$\pi(X_T) = 0, \quad P(X_t) = 1_{\{X_t \geq 0\}}$$

and the ruin probability with deficit larger(absolute value) than z

$$\pi(X_T) = 1_{\{X_T < -z\}}, \quad P(X_t) = 0$$

[13]The motivation of Karl Pearson (1924) [**?**] was to generalize a binomial distribution identity of Alexander Cuming and De Moivre to continuous distributions. In doing so, he provided a useful parametric model which includes some of the most important distributions in probability and statistics: Gaussian, exponential, Gamma, Beta, etc. The corresponding class of diffusions having these distributions as stationary measures includes also some of the most useful diffusion models, like Ornstein-Uhlenbeck, the CIR/Feller process, (quite popular in the modeling of interest rates), the Black-Scholes model, ...

Under the markovian model (46), expectations of the form (49) satisfy "morally" Kolmogorov equations:

$$\begin{aligned} G\ v_t(x) - \frac{\partial v(t,x)}{\partial t} = 0 \qquad & \text{for } u \geq 0 \\ v(t,x) = \pi(x) \ \text{ for } u \leq 0 \\ v_0(x) = P(x) \ \text{ for } u \geq 0 \end{aligned}$$

where

$$Gh(x) = [\mu(x)\,\frac{\partial}{\partial x} + d(x)\,\frac{\partial^2}{\partial x^2}]\ h(x) + \lambda \int_0^1 (h(x-z) - h(x))P(dz)$$

is the generator of the associated Markovian semigroup and $d(x) := \frac{\sigma^2(x)}{2}$.

We take now a Laplace-Carson transform in time:

$$v_q(x) = \int_0^\infty q e^{-qt} v(t,x)dt = \mathbb{E}_x\left(\ P(X_{\mathcal{E}(q)}))\ 1_{\{T \geq \mathcal{E}(q)\}} + \pi\left(X_T\right))\ 1_{\{T < \mathcal{E}(q)\}}\right) \tag{50}$$

which may also be interpreted as the ruin probability before an exponentially distributed time span $\mathcal{E}(q)$ with parameter q.

Taking into account the boundary conditions and putting $\pi_B(x) = \lambda \int_x^\infty \pi(x-u)B(du)$, we find the easier time independent problem:

$$\mathcal{G}\ v_q(x) - q\, v_q(x) + \pi_B(x) + qP(x) = 0$$

where

$$\mathcal{G}h(x) = [\mu(x)\,\frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2}\,\frac{\partial^2}{\partial x^2} - \lambda]\ h(x) + \lambda \int_0^x h(x-z)P(dz) \tag{51}$$

$$:= [g(x,D]\ h(x) + \lambda \int_0^x h(x-z)P(dz), \tag{52}$$

and $g(x,D) = \frac{\sigma^2(x)}{2}\,D^2 + \mu(x)\,D - \lambda$ denotes the differential part of our original operator.

For example, using our boundary conditions, we find that the killed ruin and survival probabilities satisfy respectively:

$$\begin{aligned} \mathcal{G}\psi_q(x) - q\psi_q(x) + \lambda\bar{P}(x) = 0 \\ \mathcal{G}\overline{\psi}_q(x) - q\overline{\psi}_q(x) + q = 0 \end{aligned}$$

Adding finally the boundary conditions, we arrive at:

$$\begin{aligned} & \mathcal{G}\ v_q(x) - qv_q(x) = -\pi_B(x) - qP(x), \ \text{ for } u \geq 0 \\ & v_q(1) = P(1) \\ & \begin{cases} g(0,D)v_q(0) - qv_q(0) = -\int_0^\infty \pi(-u)\nu(du) - qP(0) \\ v_q(0) = \pi(0_-) \qquad \text{if } \sigma_0 > 0 \end{cases} \end{aligned} \tag{53}$$

Notes: 1) The last boundary condition is just the limit of the integro-differential equation when $x \to 0$.

2) It may be convenient to separate the Gerber-Shiu function into two parts:

$$v_q(x) = v_{q,j}(x) + v_{q,d}(x)$$

which represent respectively the payoff in the case of crossing the boundary via a jump, and via smooth crossing [**?**].

The functions $v_{q,j}(x), v_{q,d}(x)$ satisfy respectively:

$$\begin{aligned}
&\mathcal{G}\ v_{q,j}(x) - q\, v_{q,j}(x) = -\pi_B(x), \quad \text{for } x \geq 0, \quad v_{q,j}(0) = 0 \\
&v_{q,j}(\imath) = 0 \\
&g(0, D)v_{q,j}(0) = -\int_0^\infty \pi(-u)\nu(du)
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{G}\ v_{q,d}(x) - q\, v_{q,d}(x) = -qP(x), \quad \text{for } x \geq 0 \\
&v_{q,d}(\imath) = P(\imath) \\
&g(0, D)v_{q,d}(0) - qv_{q,d}(0) = -qP(0) \\
&v_{q,d}(0) = \pi(0_-) \qquad \text{if } \sigma_0 > 0
\end{aligned}$$

3) In the case of phase-type claims of order m, the integro-differential term may always be removed, at the price of increasing the order of the differential part by m, as illustrated in section **??**.

4) The killed ruin probability is an q-harmonic function, but the killed survival probability is harmonic only at $q = 0$ (in which case it differs from the corresponding harmonic function called scale only by a proportionality constant.

5) Penalty-reward functions are known as Gerber-Shiu functions in the actuarial literature.

6) In general, Feynman-Kac equations of the type encountered above may be obtained formally from a "dictionary" associating Markov processes to their generator operators. More precisely, we are effectively using a table whose rows contain fixed formulas like $Gf(x) = 0, Gf(x) - qf(x) = 0, Gf(x) + 1 = 0$, etc (corresponding to particular types of problems), and whose columns contain generators of semigroups. These equations are easiest to establish for discrete time random walks, when they reduce to recurrences (like (**??**) below). The continuous time case encountered here (while just a limit of the discrete time case) involves several technical issues like conditions to ensure the existence of solutions in a classical sense rather than in the sense of generalized functions, ...

Solution of the IDE by Laplace transform. We illustrate now the method of solution of these IDE by Laplace transform. This approach works even in the case of polynomial state-dependent coefficients, producing differential equations for the Laplace transform – see section 9.2, but is most efficient in the case of Levy models.

9.2 The double Laplace transform of Gerber-Shiu functions

In the classic state independent spectrally negative Levy case (with constant coefficients), a double Laplace transform in time and the initial condition yields:

$$v_q^*(\theta) = \int_0^\infty e^{-\theta x} v_q(x) dx = \tag{54}$$

$$\frac{q(P^*(\theta) - P^*(q_+)) + \int_0^\infty \pi(-x)(\nu_x^*(\theta) - \nu_x^*(q_+))dx - \frac{\sigma^2}{2}\pi(0)(\theta - q_+)}{q - \kappa(\theta)} \tag{55}$$

where $\nu_x^*(\theta) = \int_X^\infty e^{-\theta(y-x)}\nu(dy)$, $P^*(\theta) = \int_0^\infty e^{-\theta x}P(x)dx$ and $q_+ = \Phi(q)$ is the largest and unique non-negative root of the equation $\kappa(\theta) = q$ [**?**].

In the survival case for example, this yields:

$$\overline{\psi}_q^*(\theta) = \int_0^\infty e^{-\theta x}\overline{\psi}_q(x)dx = \frac{q(\theta^{-1} - q_+^{-1})}{q - \kappa(\theta)}$$

We turn now to the double Laplace transform of Gerber-Shiu functions (scale, ruin and survival probabilities, ...) for Wong-Pearson jump-diffusions.

Lemma 9.1 *Under assumption (S), the double Laplace-Carson transform*

$$v_q^*(s) = \int_0^\infty e^{-sx}\, v_q(x)dx = \int_{x=0}^\infty e^{-sx} \int_0^\infty qe^{-qt}v(t,x)dt\, dx \tag{56}$$

satisfies for $s > 0$ the linear ODE:

$$\begin{aligned} &d_2 s^2\, v_q^{*\prime\prime}(s) + \left(-rs\ - d_1 s^2 + 4d_2 s\right)\, v_q^{*\prime}(s) + (\kappa(s) - q - 2d_1 s - r + 2d_2)\, v_q^*(s) \\ &- v_q(0)\,(c + d_0 s - d_1) - d_0\, v_q'(0) + \pi_B^*(s) + qP^*(s) = 0 \end{aligned} \tag{57}$$

where $\pi_B^*(s) = \int_0^\infty e^{-sx}\,\pi_B(x)\ dx$ and $P^*(s) = \int_0^\infty e^{-sx}\,P(x)\ dx$, and $\kappa(s) = cs + d_0 s^2 + \lambda(b^*(s) - 1)$.

Proof: This follows by taking the Laplace transform in x of the IDE:

$$\begin{aligned} &[(p + rx)\, D + (\sum_{i=0}^{2} d_i x^i)\, D^2 - (\lambda + q)]v_q(x) + \lambda \int_0^x v_q(x-u)b(u)du \\ &+\pi_B(x) + qP(x) = 0 \text{ for } x \geq 0 \\ &v_q(0) = \pi(0_-) \qquad \text{if } \sigma_0 > 0. \end{aligned} \tag{58}$$

Notes: 1) It turns out that when $\pi_B^*(s) + qP^*(s) \neq 0$, then the analiticity of $v_q^*(s)$ for $s \geq 0$ allows determining $v_q(0)$ and $v_q'(0)$ and solving finally for $v_q^*(s)$; for the Levy case, we obtain in this way the equation (55 – see [**?**]. For the non-Levy affine case, see section **??** below.

9.3 The scale function and two-sided exit problems

Interestingly, the answer to several two-sided passage problems for spectrally negative processes may be expressed as quotients involving the so called scale function $W(x)$ which is proportional to the survival probability $1-\psi(x)$ and its "killed/shifted version" $W^{(q)}(x)$. Up to a constant, $W^{(q)}(x)$ is defined as the solution (in a sense "adapted to the available regularity") of the homogeneous Dirichlet equation:

$$\begin{aligned} GW_q(x) - qW_q(x) &= 0 \\ W_q(x) &= 0, \quad x<0 \end{aligned} \tag{59}$$

Note: The space of solutions of these equations for generalized Wong-Pearson processes is of dimension one. Indeed, putting $\pi_B^*(s) = P^*(s) = 0$ in the Laplace transform of Lemma 9.1, we arrive at:

$$\begin{aligned} &d_2 s^2 \hat{W}_q^{*\prime\prime}(s) + \left(-rs - d_1 s^2 + 4d_2 s\right) \hat{W}_q^{*\prime}(s) + \left(\kappa(s) - q - 2d_1 s - r + 2d_2\right) \hat{W}_q^*(s) \\ &- W_q(0)\left(c + d_0 s - d_1\right) - d_0\, W_q'(0) = 0 \end{aligned} \tag{60}$$

When $d_0 = 0$, this leaves one free parameter $W_q(0)$, and when $d_0 \neq 0$, the presence of Brownian motion ensures $W_q(0) = 0$, leaving again one free parameter $W_q'(0)$.

Exemple 9.1 *For Levy models, the equation (60) becomes*

$$(\kappa(s) - q)\hat{W}_q(s) = (sd_0 + c))\, W_q(0) + d_0 W_q^{'}(0) \tag{61}$$

The two cases may be unified here by defining W_q to be the unique continuous increasing function $W_q : [0,\infty) \to [0,\infty)$ with Laplace transform

$$\int_0^\infty e^{-\theta x} W_q(x)\, dx = \left(\kappa(\theta) - q\right)^{-1}, \qquad \theta > \Phi(q)$$

This imposes in the two cases the jump discontinuities $W_q(0) = \frac{1}{\mu}$ and $W_q'(0) = \frac{1}{d_0}$, respectively.

Notes: *1) As may be seen from the fluctuation theory of spectrally negative Lévy processes (e.g. Bertoin (1996), Bingham (1975)), the scale and killed survival probability are related for $x \geq 0$ via*

$$\overline{\Psi}_q(x) = \Phi(q)^{-1} W_q(x) - \overline{W_q}(x), \tag{62}$$

where $\overline{\Psi}_q(x) = \frac{1-\psi_q}{q}$ denotes the Laplace transform in t of the finite time survival probabilities.

*2) The existence of W_q, i.e. the fact that $\left(\kappa(\theta) - q\right)^{-1}$ is indeed a Laplace transform for $\theta > \Phi(q)$, is verified in [**?**]. Alternatively, one may start with $q = 0$, in which case we have the Laplace transform of the infimum, and use subsequently a change of measure.*

Exemple 9.2 *In the case of the classical risk model $X_t = ct - S_t$ where S_t is compound Poisson with exponential "claims" of parameter μ (which is of phase-type with $\boldsymbol{\beta} = 1$, $\mathbf{B} = -\mu$), the Laplace exponent is $\kappa(\theta) = p\theta + \lambda(\frac{\mu}{\theta+\mu} - 1) = \theta(p - \frac{\lambda}{\theta+\mu})$, and we may find the scale function by decomposing $\frac{1}{\kappa(\theta)-q}$ in partial fractions.*

With exponential claims, we find:

$$\frac{1}{\kappa(\theta) - q} = \frac{\theta + \mu}{c\theta(\mu + \theta) - \lambda\theta - q(\mu + \theta)} = \frac{\theta + \mu}{c\theta^2 + \theta(c\mu - \lambda - q) - q\mu} = \frac{s + \mu}{c(s - q_+)(s - q_-)}$$

where $q_\pm$ denote the nonnegative/negative roots of the Cramer Lundberg equation

$$\kappa(\theta) - q = 0 \Longleftrightarrow c\theta^2 + \theta(c\mu - \lambda - q) - q\mu = 0$$

The partial fractions decomposition is:

$$\frac{1}{\kappa(\theta) - q} = \frac{\theta + \mu}{c(\theta - q_+)(\theta - q_-)} = \frac{A_+}{c(s - q_+)} - \frac{A_-}{c(s - q_-)}$$

where

$$A_\pm = \frac{c\mu + q_\pm}{q_+ - q_-}$$

and finally

$$W^{(q)}(x) = c^{-1}(A_+ e^{q_+ x} - A_- e^{q_- x})$$

Note that when $q = 0$ this simplifies to

$$W(x) = \frac{1}{c(c\mu - \lambda)}(c\mu - \lambda e^{(\lambda - c\mu)x}) = \frac{1}{c - \lambda/\mu}(1 - \frac{\lambda}{c\mu} e^{(\lambda - c\mu)x})$$

which is proportional to the survival probability $\overline{\psi}(x) = 1 - \frac{\lambda}{c\mu}\mathrm{e}^{(\lambda - c\mu)x}$.

Exemple 9.3 *Consider now the perturbed risk model with exponential claims, i.e. $r = 0, d_1 = d_2 = 0, q = 0$. Now, (70) becomes:*

$$\begin{aligned} &[d_0 D^3 + (c + \mu d_0)\, D^2 + (\mu c - \lambda) D]\psi(x) = 0 \\ &d_0\psi''(0) + c\psi'(0) - \lambda\psi(0) = -\lambda, \quad \psi_q(1) = 0, \end{aligned}$$

with the solution

$$\psi(x) = \psi_s(x) + \psi_d(x)$$

where all these three functions are of the form $\sum_i C_i \mathrm{e}^{-\gamma_i x}$, γ_i being the roots of $\gamma^2 - (\tilde{\mu} + \tilde{c})\gamma + (\tilde{c}\mu - \tilde{\lambda}) = 0$, with $\tilde{c} = c/d_0, \tilde{\lambda} = \lambda/d_0$.

To satisfy $\psi_s(0) = 0, d_0\psi''(0) + c\psi'(0) = \lambda$, we must take $C_{i,s} = \frac{k}{\mu(\gamma_i - \gamma_{3-i})}, i = 1, 2$, where $k = (\gamma_1 - \mu)(\gamma_2 - \mu)$. The conditions $\psi_s(0) = 1, d_0\psi''(0) + c\psi'(0) = 0$ lead to $C_{i,d} = \frac{\gamma_i - \mu}{\gamma_i - \gamma_{3-i}}$.

.....

In the affine case, we can solve explicitly the first order ODE for the Laplace transform:

Théorème 9.1 *When $d_2 = 0$, for any $r > 0, c > 0, s > 0$ and assuming $c > d_1$ in the particular case $d_0 = 0, d_1 > 0$, it follows that the solutions of the equation (59) form a one dimensional space, and are determined up to a constant by having the Laplace transform $\hat{W}_q(s) = \int_0^\infty e^{-sx} W_q(x)dx$ satisfying*

$$\begin{aligned} &\left(-rs - d_1 s^2\right) \hat{W}_q^{*\prime}(s) + (\kappa(s) - q - 2d_1 s - r)\, \hat{W}_q^*(s) \\ &- \hat{W}_q(0)\,(c + d_0 s - d_1) - d_0\, \hat{W}_q'(0) = 0, \quad W_q(0) = 0 \text{ if } d_0 > 0 \end{aligned} \tag{63}$$

and

$$\begin{aligned} &\hat{W}_q(s) + W_q(0) \\ &= \int_s^\infty \left(\frac{W_q(0)\,(p + d_0 x - d_1) + d_0\, W_q'(0)}{B(x)} \right) E_q(s,x)dx, \quad W_q(0) = 0 \text{ if } d_0 > 0. \end{aligned} \tag{64}$$

and

$$\lim_{s\to\infty} \hat{W}_q^*(s) = 0$$

Finally ....

Exemple 9.4 *For the "generalized OU" process of Hadjiev with $r = 1$, the Laplace transform is:*

$$\begin{aligned} &\hat{W}_q(s) + W_q(0) \\ &= \int_s^\infty \left(W_q(0)\,(p + d_0 x) + d_0\, W_q'(0)\right) \frac{m_q(x)}{m_q(s)}\; e^{\lambda \int_s^x \overline{F}^*(z)dz} dx \end{aligned} \tag{65}$$

Exemple 9.5 *When $d_0 = 0, d_1 > 0, r > 0$ and with exponential jumps ....*

9.4 Examples of the ODE approach

9.4.1 Generalized Pearson process with exponential claims

Following the pioneering work of Segerdahl (1942), the case of exponential claims with rate μ has been studied by several authors like [**?**], [**?**], [**?**]. The finite time ruin probability $\psi(t,x)$ satisfies:

$$\begin{aligned} &[d(x)\, D^2 + c(x)\, D - \lambda]\psi(t,x) + \lambda\mu \int_0^x \psi(t,u) e^{-\mu(x-u)} du + \lambda e^{-\mu x} \\ &= \frac{\partial}{\partial\, t}\psi(t,x) \text{ for } x \geq 0 \\ &\psi(t,0) = 1 \qquad \text{if } \sigma(0) > 0, \quad \psi(t,\infty) = 0 \\ &\psi(0,x) = 1_{x<0} \end{aligned} \tag{66}$$

where D denotes derivative with respect to x.

The killed ruin probability satisfies:

$$[d(x)\,D^2 + c(x)\,D - (\lambda+q)]\psi_q(x) + \lambda\mu\int_0^x \psi_q(u)e^{-\mu(x-u)}du$$
$$= -\lambda e^{-\mu x} \text{ for } x \geq 0 \tag{67}$$
$$[pD + d_0D^2 - \lambda - q]\psi_q(0) = \lambda, \quad \psi_q(\infty) = 0$$
$$\psi_q(0) = 1 \qquad \text{if } \sigma(0) > 0 \tag{68}$$

The integro-differential convolution term part may be easily removed. Indeed, multiplying with $e^{\mu x}$ and differentiating the equation yields the ODE:

$$D[e^{\mu x}[d(x)\,D^2 + c(x)\,D - (\lambda+q)]]\psi_\mu(x) + \lambda\mu\psi_q(x)e^{\mu x}$$
$$= 0 \quad \text{for } x \geq 0 \tag{69}$$
$$\psi_q(\imath) = 0, \quad \psi_q(0) = 1 \quad \text{if } v(0) > 0$$

Using now the relation $e^{-\mu x}D(e^{\mu x}f) = (\mu + D)f$, we arrive at the ODE:

$$[(\mu+D)(d(x)\,D^2 + c(x)\,D - \lambda - q) + \lambda\mu]\psi_q(x) =$$
$$[(\mu+D)(d(x)\,D^2 + c(x)\,D - q) - \lambda D]\psi_q(x) = 0 \text{ for } x \geq 0$$

and finally

$$[d(x)\,D^3 + (d'(x) + \mu d(x) + c(x))\,D^2 + (\mu c(x) + c'(x) - q - \lambda)D - \mu q]\psi_q(x) = 0$$
$$[d_0D^2 + cD - (\lambda+q)]\psi_q(0) = \lambda + q, \quad \psi_q(\imath) = 0,$$
$$\psi_q(0) = 1 \qquad \text{if } v(0) > 0$$

Notes: 1) This manipulations amount to applying to the equation the operator $\mu + D$, which is formally a left inverse of the convolution operator with the exponential kernel [14].

2) When $d(x) = 0$ (and also, effectively, when $q = 0$), the equation is of second order; its solution is therefore a classical hypergeometric series.

The Cramer-Lundberg case: When $r = 0, d(x) = 0$, (70) becomes:

$$[c\,D^2 + (\mu c - q - \lambda)D - \mu q]\psi_q(x) = 0$$
$$c\psi_q'(0) - (\lambda+q)\psi_q(0) = \lambda + q, \quad \psi_q(\imath) = 0,$$
$$\psi_q(0) = 1 \qquad \text{if } v(0) > 0$$

[14] Alternatively, a substitution $\psi_q(x) = e^{-\mu x}f_q(x)$ followed by a change of variables $x' = \mu x$ changes the equation into an ODE with respect to $F_q(x) = \int_0^x f_q(u)du$:

$$[d_\mu(x)\,D(D^2 - 2D + 1) + p_\mu(x)\,D(D-1) - (\lambda+q)D + \lambda]F_q(x) + \lambda = 0$$
$$= 0 \text{ for } x \geq 0 \tag{70}$$
$$F'_q(0) = 1 \quad \text{if } \sigma(0) > 0, \quad F'_q(\infty) = 0, \quad p\mu F''_q(0) + d_0\mu^2 F'''_q(0) = q$$

where $p_\mu(x) = p\mu + rx, d_\mu(x) = d_0\mu^2 + d_1\mu x + d_2x^2$.

with the solution

The Segerdahl/Paulsen result. Putting $\hat{\lambda} := \frac{\lambda}{r}, \hat{p} = \frac{p}{r}, \hat{q} := \frac{q}{r}$, the equation (70) becomes:

$$(\hat{p}+x))\, D^2 + (\mu(\hat{p}+x) + 1 - \hat{q} - \hat{\lambda})D - \mu\hat{q}]\psi_q(x) = 0, \quad \psi_q(1) = 0$$

and its solution is:

$$\psi_q(x) = \left(\frac{\lambda}{p\mu}\right) \frac{\int_x^\infty e^{-\mu s}(1+x/\hat{p})^{\hat{\lambda}-1}(s-x)^{\hat{q}} ds}{U\left(1+\hat{q}, 2+\hat{q}+\hat{\lambda}, \mu\hat{p}\right)} \tag{71}$$

$$= \left(\frac{\lambda}{p\mu}\right) \frac{e^{-\mu x}(1+x/\hat{p})^{\hat{q}+\hat{\lambda}} U\left(1+\hat{q}, 1+\hat{q}+\hat{\lambda}, \mu(x+\hat{p})\right)}{U\left(1+\hat{q}, 2+\hat{q}+\hat{\lambda}, \mu\hat{p}\right)} \tag{72}$$

where

$$U(a,b,y) = \int_0^\infty t^{a-1}\, e^{-yt}\, (1+t)^{b-a-1}\, dt \tag{73}$$

is the confluent Tricomi hypergeometric function which satisfies (see expression 13.4.7 in Abramowitz and Stegun(1970))

$$U\left(1+\hat{q}, 1+\hat{q}+\hat{\lambda}, \mu\hat{p}\right) + (1+\hat{q})\, U\left(2+\hat{q}, 2+\hat{q}+\hat{\lambda}, \mu\hat{p}\right) = U\left(1+\hat{q}, 2+\hat{q}+\hat{\lambda}, \mu\hat{p}\right) \tag{74}$$

Notes: 1) The Tricomi hypergeometric function (73) reduces to a Gamma function when $b = a+1$; it is also the Laplace transform of the inverted Beta distribution.

2) For Segerdahl's problem with $q = 0$, using $U(1,b,z) = e^z z^{1-b}\Gamma(b-1,z)$, this answer reduces to the classical:

$$\psi(x) = \frac{\hat{\lambda}\, \Gamma\left(\hat{\lambda}, \mu(x+\hat{p})\right)}{\Gamma\left(\hat{\lambda}+1, \mu\,\hat{p}\right)}$$

3) The identity (74) allows simplifying the normalization constant in Paulsen's Example 2.1 [**?**]. For example, Segerdahl's result is written there in the equivalent form:

$$\psi(x) = \frac{\hat{\lambda}\Gamma\left(\hat{\lambda}, \mu(x+\hat{p})\right)}{e^{-\mu\hat{p}}(\mu\hat{p})^{\hat{\lambda}} + \hat{\lambda}\Gamma\left(\hat{\lambda}, \mu\,\hat{p}\right)}$$

The current simplification appears in [**?**].

4) The double Laplace transforms with respect to time and space may also be obtained explicitly – see [**?**] or section 9.2 for the affine case. The intervening expressions (21-22) in section 9.2 have probably interesting probabilistic interpretations ....

9.5 First passage for Pearson diffusions

Examples:

1. For OU process centered at 0, with killing q and $\Gamma = \frac{\sigma^2}{2}D^2 - \lambda x D$, the monotone harmonic functions are: $\varphi_\pm(x) = H_{-q/\lambda}(\pi = (\sigma\sigma')^{-1}(\mu - r)x\sqrt{\lambda}/\sigma)$ where $H_q(x)$ is the Hermite function defined as the decreasing solution of the normalized generator $\frac{1}{2}H_q''(x) - xH_q'(x) - qH_q(x) = 0$.

2. For Bessel process of order ν, with $\Gamma = \frac{1}{2}D^2 + \frac{2\nu+1}{2x}D$ (which is the law of the modulus of Brownian motion in dimension $n = 2\nu + 2$), the monotone harmonic functions are: $x^{-\nu}I_\nu(x\sqrt{2q})$ and $x^{-\nu}K_\nu(x\sqrt{2q})$ where I_ν, K_ν are the modified Bessel functions defined as monotonic solutions of: $\frac{1}{2}x^2 I_\nu''(x) + xI_\nu'(x) - (x^2 - \nu^2)I_\nu(x) = 0$.

2) For the processes in the last two examples, the the scale functions are given respectively by $S(q/\lambda, x/\sigma, a/\sigma)$ and $S(\nu, x\sqrt{2q}, a\sqrt{2q})$ where $S(q, x, y)$, the parabolic cylinder function, satisfies the Weber equation

$$Z''(x) - xZ'(x) - qZ(x)$$

– see [**?**] (note that probabilistically, this is the quasi-stationary distribution of $[x, y]$ for a transient OU process with variance 2, killed at rate q???).

10 Transient distributions and first-passage probabilities of Markovian processes

Markovian semigroup transition probabilities obey, depending on the state-space of the process under consideration, Chapman-Kolmogorov partial (integro-) differential (difference) equations. Their resolution presents numerical challenges even in the simplest finite state-space case, which amounts to the computation of the matrix-exponential function e^{tG} – see for example the classic Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later, (2003) by Cleve Moler and Charles Van Loan.

In the simplest model with countable infinite state-space, the transitions of the M/M/1 queue from 0 to n, the simplest method seems to be obtaining and inverting the Laplace transform. This method is also the favorite for the "ruin problem" in actuarial science. Note however that inversion leads already to hypergeometric functions even in the simplest examples.

Finite-time ruin probabilities obey the same differential equations as the transition probabilities (with different boundary conditions) and are in fact "a notch harder". The only known explicit analytical results for now are for Brownian motion and for the Cramer-Lundberg process with exponential claims ([**?**]) (the "Kou process", i.e. Brownian motion plus exponential claims, is also feasible probably).

For the Cramer-Lundberg process with exponential claims, Bessel functions intervene already (not surprisingly, since the semigroup involves the hypergeometric function ${}_0F_1(\binom{-}{2}; \lambda\mu t x)$, which is related to the Bessel function of order 1 with imaginary argument).

The difficulties encountered in this cases with explicit inversion via Bromwich's formula –see [**?**], [**?**], [**?**]– do not seem to recommend the "analytic approach" (i.e explicit determination of the spectrum and integration in the complex plane).

The inversion of the Laplace transform may be accomplished via the Gaver-Stehfest algorithm or via Post-Widder inversion, also known as Erlangization randomization (approximating a finite-time s by an Erlang distributed random variable with the same mean). Note that this approach is particularly attractive in the case of phase-type claims, since it allows combining the phase-type structure of the claims with that of the Erlang horizon.

However, with a continuous state-space and state dependent coefficients, as is often the case in mathematical finance, obtaining the transform explicitly seems unnecessary, and Erlang/Laguere series expansions –see attached joint project? – may be preferrable (the coefficients of these series may be viewed as projections of the Laplace transform along certain bases).

The crux of the problem seems therefore the form under which the Laplace transform is represented (as a function, or as coefficients, ...) and the inversion of the transform.

Since all approaches end up with expansions, asymptotic considerations (the diffusion or the Arfwedson-Hoglund approximation) should also be taken into account.

A new very promising approach recently proposed by Picard and Lefevre [**?**] and Ignatov and Kaishev [**?**] is based on the obvious probabilistic decomposition:

$$\overline{\psi}(t,x) = \mathrm{e}^{-\lambda t} \sum_{n=0}^{\infty} \int_{0 \le s_1 \le ... \le s_n} F_S(ds_1, ..., ds_n) P[T_{(n)} \le t, T_{(1)} \ge v_1, ..., T_{(n)} \ge v_n \qquad (75)$$

where $v_k = [h^{-1}(s_k)]_+$ is the first possible time when the arrival of the k'th claim leading to a cumulative loss of s_k will not cause ruin, given a monotone premium function $h(t)$ with $h(0) = x$.

We recall now the implementation in the simplest case of discrete size claims, when one may take advantage of the recursive construction of Appell/Sheffer expansions.

Consider a compound-Poisson distribution $P_n(s) = Pr[\sum_{i=1}^{N(s)} Z_i = n]$ where $Z_i \in \{1, 2, ...\}$ are i.i.d. and $N(s)$ is an independent Poisson process of rate λ, and put $P_n(t) := \mathrm{e}^{-\lambda s} p_n(\lambda s), s \in \mathbb{R}^+$. It is easy to see that $p_n(s)$ are polynomials satisfying the convolution property, and therefore quite easily computable recursively, in terms of the sequence $g = (g_1, g_2, ...)$, where $g_i := Pr[Z_1 = i], i = 1, 2, ...$ In other words, the solution at time s of the Chapman-Kolmogorov equations for $P_n(s)$ corresponding to convolution with g and boundary conditions $P_n(0) = \delta_0(n)$, may be efficiently obtained.

As seen from the work of Picard and Lefevre [**?**] and Ignatov and Kaishev [**?**], the first passage problem is only "one notch" harder than this semigroup computation, in the sense that the survival probabilities conditioned on $n-$ claims satisfy the same convolution equations and equivalent Appell recursions (for the Poisson process for example, this is just $\bar{\psi}_n'(s) = \bar{\psi}_{n-1}(s)$) but with different boundary conditions. More precisely, the distribution $\bar{\psi}_n(s)$ of the number of claims before time s of a "non-ruined" Cramer-Lundberg risk model with n claims satisfy

$$\bar{\psi}_n(s) = 0 \; \forall s \le v_n$$

with some easily computable boundary conditions v_n.

It has been long known that the "Sheffer polynomials" associated to such a shift in the boundary conditions satisfy simple recursive relations; it follows thus that finite-time survival probabilities may be computed just as efficiently as discrete compound distributions!

11 Sheffer and convolution polynomials

11.1 Polynomials of convolution and binomial type

Définition 11.1 *A sequence of polynomials $p_n(x)$ is of binomial type if it satisfies the identity:*

$$p_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} p_k(x) p_{n-k}(y).$$

A sequence of polynomials $q_n(x)$ is of convolution type (or divided differences) if it satisfies the identity:

$$q_n(x+y) = \sum_{k=0}^{n} q_k(x) q_{n-k}(y). \tag{76}$$

It is easy to check that p_n is of binomial type iff $q_n = \frac{p_n}{n!}$ is of convolution type and that if $p_0 \neq 0$, than a sequence of binomial type satisfies: a) $p_0(x) = 1 \ \forall x, \ b) p_n(0) = 0 \forall n \geq 1$, and c) the degree of p_n and q_n cannot exceed n. In the case of equality, binomial sequences are also called **basic**.

The classic example of a pair of a binomial and a convolution type sequence is provided by $p_n = x^n$ and $q_n = \frac{x^n}{n!}$. Other examples:

1. The sequence of "lower factorials" defined by:
$$(x)_n = x(x-1)(x-2)...(x-n+1)$$
2. The "upper factorials" defined by:
$$(x)^n = x(x+1)(x+2)...(x+n-1)$$
3. The Abel polynomials defined by:
$$a_n(x) = x(x-an)^{n-1}$$
4. Laguerre polynomials
5. The Touchard polynomials defined for $\lambda > 0$ by $p_n(\lambda) = E(X^n)$, X is a Poisson random variable with expected value λ.

By definition, all these polynomials satisfy convolution identities. For example, the normalized lower factorials satisfy:

$$\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}$$

11.1.1 Characterization by the generating function/sequence

It is easy to show that polynomial sequences of binomial (convolution) type are precisely those whose generating functions $p(x,z) = \sum_n p_n(x)z^n/n!$ ($q(x,z) = \sum_n q_n(x)z^n$) are of the form

$$\sum_n q_n(x)z^n = \mathrm{e}^{xg(z)} \tag{77}$$

for some function satisfying $g(0) = 0, g'(0) \neq 0$ (this follows from the multiplicative property $q(x+y,z) = q(x,z)q(y,z)$).

Définition 11.2 *The sequence g_n defined by $g(z) = \sum_{n=1} g_n z^n$ is called the* **coefficients generating sequence** *(since it yields easily the coefficients of the binomial polynomial sequences).*

Note: When the coefficients generating sequence is nonnegative and sums to one, then $g(z) = \sum_{n=1} g_n z^n$ has a probabilistic interpretation, i.e. $g(z) = Ez^Z$ is the generating function of a random variable with values in $\mathbb{N}_+$. In this case, the convolution polynomials $q_n(\lambda), \lambda > 0$ coincide up to the normalization $e^{-\lambda}$ with the distribution of the compound sum $S = \sum_{i=1}^{N_\lambda} Z_i$, where Z_i are i.i.d. with the same distribution as Z, and N_l is a Poisson random variable with expected value λ.

One efficient way to compute compound Poisson convolution distributions $q_n(x)$ is the recursion

$$q_n(x) = \sum_{i=1}^{n} q_{n-i}(x) g_i \frac{\mu\, i}{n} \tag{78}$$

where μ is the Poisson parameter (and Panjer's recursion generalizes this to the case of negative binomial compounded distributions).

Note: For any Levy process $S(t)$ (in particular for compound Poisson processes), the moments are also of binomial type (while the distributions are of convolution type), due to the multiplicativity of the moment generating function. For example, for the Poisson process of rate 1, the moments are the upper factorials $m_n(t) = (t)^n$.

11.1.2 Characterization by Δ/lowering operators

A sequence of polynomials $q_n(x)$ of convolution type may be characterized by the associated linear operator defined on the basis $q_n(x)$ by $Qq_n = q_{n-1}$.

It may be shown that for any sequence of polynomials $q_n(x)$ of convolution type, the associated linear operator defined by $Qq_n = q_{n-1}$ commutes with the derivative operator D and the shift operators $T_h f(x) := f(x+h)$.

Example: The usual powers $p_n(x) = x^n, q_n(x) = \frac{x^n}{n!}$, have as associated operator the derivative $Q = D = \frac{d}{dx}$, i.e.

$$Dq_n(x) = q_{n-1}(x) \Longleftrightarrow Dp_n(x) = np_{n-1}(x),$$

and the operators associated with lower/upper factorials are respectively $Q = T - I$ and $Q = I - T^{-1}$, where $Tf(x) := f(x+1)$.

Définition 11.3 *A Δ/lowering operator is an operator Q acting on polynomials which commutes with the shift operators $T_h f(x) = f(x+h)$ (and hence with the derivative operator D), and such that Qx is a constant.*

It may be shown that the correspondence between basic convolution (or binomial) sequences and Δ/lowering operators is one to one!

Formally, Δ/lowering operators may be represented as $Q = h(D)$ for some formal (not necessarily convergent) power series $h(x)$. Thus, the upper factorial operator satisfies formally $Q = I - T^{-1} = I - e^{-D}$ (since $T_h = e^{hD}$ is true formally, coinciding with the MacLaurin expansion).

The lowering operator satisfies formally the identity

$$D = g(Q) = \sum_{n=0}^{\infty} g_n Q^n,$$

where g_n is the coefficients generating sequence, and thus $Q = h(D)$ where $h = g^{-1}$ is the compositional inverse of the function $g(z)$ appearing in the generating function. It also holds that Q is of the form $Q = DU^{-1}$, where $U = u(D)$ is an invertible operator, in the sense that q_n satisfy a recurrence

$$q_n'(x) = u(D)q_{n-1}(x).$$

The convolution sequence may be recovered then via

$$n!q_n(x) = (xU^n)(x^{n-1})$$

Exemple 11.1 *For the usual powers $p_n(x) = x^n, q_n(x) = \frac{x^n}{n!}$, we have $g(x) = x, h(x) = x$, the associated operator is simply $Q = D$, and $U = Id$.*

Note the representation of powers via a multiple integral:

$$\frac{x^n}{n!} = \int_{y_n=0}^{x} dy_n \int_{y_{n-1}=0}^{y_n} dy_{n-1} ... \int_{y_1=0}^{y_2} dy_1, n \geq 1 \tag{79}$$

Exemple 11.2 *The operator and polynomials corresponding to geometric summands are:*

The other classic convolution polynomials do not correspond to compound Poisson distributions. For example,

Exemple 11.3 *The Abel polynomials $x(x-an)^{n-1}/n!$ correspond to the operator $Q = DT_a$, i.e. satisfy*

$$a_n'(x) = a_{n-1}(x-a)$$

The coefficients generating sequence is

Exemple 11.4 *The Laguerre polynomials correspond to the operator $Q = D(D-I)^{-1}$, i.e. satisfy*

$$l_n'(x) = l_{n-1}(x)' - l_{n-1}(x)$$

The coefficients generating sequence is

Exemple 11.5 *The upper factorials polynomials have coefficients generating sequence*

11.2 Sheffer polynomials

Let $v = (v_1 \leq v_2 ... \leq v_n ...)$ denote an infinite sequence of real numbers.

Consider the "Apell polynomials" $\bar{A}_n(x)$ represented by shifted multiple integrals:

$$\bar{A}_n(x) = \int_{y_n=v_n}^{x} dy_n \int_{y_{n-1}=v_{n-1}}^{y_n} dy_{n-1} ... \int_{y_1=v_1}^{y_2} dy_1, n \geq 1 \tag{80}$$

and satisfying the boundary conditions $\bar{A}_1(v_1) = 0, \bar{A}_1(x) > 0, x \geq v_1, ... \bar{A}_n(v_n) = 0, \bar{A}_n(x) > 0, x \geq v_n, ...$

For example, the first Apell polynomials are:

$$\bar{A}_1(x) = x - v_1, \bar{A}_2(x) = (x - v_2)(x/2 - v_1 + v_2/2), ...$$

It turns out that the Apell polynomials satisfy a recurrence

$$\bar{A}_n(x + y) = \sum_{k=0}^{n} \frac{x^k}{k!} \bar{A}_{n-k}(y) \tag{81}$$

similar to the convolution identity satisfied by their "basic" sequence $q_k(x) = \frac{x^k}{k!}$, and which allows their efficient computation.

Note: We reserve the notation $A_n(x) := n! \bar{A}_n(x)$ for the monic Appell polynomials.

More generally, to each convolution-type sequence $q_n(t)$ (or to each lowering operator Q), we may associate families of "Sheffer polynomials" $s_n(x)$ defined by the recursion:

$$s_n(x + y) = \sum_{k=0}^{n} s_k(x) q_{n-k}(y) \tag{82}$$

and boundary conditions $s_n(v_n) = 0$. Recursive formulas for computing Sheffer polynomials are obtained by plugging $x = 0$ in (82), which yields the decomposition formula

$$s_n(y) = \sum_{k=0}^{n} s_k(0) q_{n-k}(y) \tag{83}$$

and

Let v_n denote the largest root of $s_n(y)$ (it turns out –see [**?**]– that these polynomials have only real roots). Further substituting the largest root v_n in (83), we find a recurrence for the free coefficients:

$$s_n(0) = -\sum_{k=1}^{n} s_k(0) q_{n-k}(v_n) \tag{84}$$

11.3 Appell polynomials

The Appell polynomials defined by (80) are precisely the Sheffer polynomials obtained when $Q = D$. They may also be defined by $A_n(x) = E(x + X)^n$, where X is an arbitrary random variable. In terms of cumulants, the first monic Appell polynomials are: